



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 164752

TO: Phuong Bui
Location: rem/2A15/2C18
Art Unit: 1638
Saturday, September 10, 2005

Case Serial Number: 10/829432

From: Mary Jane Ruhl
Location: Biotech-Chem Library
Remsen 1-A-62
Phone: 571-272-2524

maryjane.ruhl@uspto.gov

Search Notes

Examiner Bui,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl
Technical Information Specialist
STIC
Remsen 1-A-62
Ext. 22524



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2005, 09:19:19 ; Search time 5327 Seconds
(without alignments)
11070.020 Million cell updates/sec

Title: US-10-829-432-3
Perfect score: 1217
Sequence: 1 ggcgcgtttcattcatca.....aaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.com.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.ste.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	359.6	29.5	928	BT009518	BT009518 Triticum
2	333.2	27.4	1434	AK071285	AK071285 Oryza sat
3	311	25.6	970	AF178976	AF178976 Zea mays
4	309	25.4	1086	AF044285	AF044285 Catharant
5	296.4	24.4	1068	AY085264	AY085264 Arabidops
6	296.4	24.4	1077	AT005238	U05238 Arabidops
7	296.4	24.4	1143	AY054287	AY054287 Arabidops
8	296.4	24.4	1185	ATCLAPS8	X75782 A.thaliana
9	294.4	24.2	831	CO805858	CO805858 Sequence
10	294.4	24.2	831	AX412633	AX412633 Sequence
11	294.4	24.2	831	AX506015	AX506015 Sequence
12	294.4	24.2	831	AY132010	AY132010 Arabidops
13	288.4	23.7	964	BT005193	BT005193 Arabidops
14	288.4	23.7	1133	AY085031	AY085031 Arabidops
15	288.4	23.7	1143	BT003977	BT003977 Arabidops
16	263.6	21.7	882	AY097421	AY097421 Arabidops
17	263.6	21.7	920	BD248389	BD248389 Gene part
18	263.6	21.7	1243	AF462823	AF462823 Arabidops
19	263.6	21.7	1311	AF043351	AF043351 Arabidops

20 172.6 14.2 300150 1 AP004598 AP004598 Oceanobac
21 170.2 14.0 11131 1 AE007523 AE007523 Clostridi
22 168.6 13.9 11037 1 AE013685 AE013685 Versinia
23 168.6 13.9 110000 1 BX936398_09 Continuation (10 o
24 168.6 13.9 220050 1 AJ414156 AJ414156 Versinia
25 168.6 13.9 290803 1 AE017128 AE017128 Versinia
26 165.2 13.6 349742 1 BX572090 BX572090 Prochloro
27 163.4 13.4 51860 8 AB013390 AB013390 Arabidops
28 162.8 13.4 521 8 BT009596 BT009596 Triticum
29 161.4 13.3 53533 1 BSU09476 BSU09476 Bacillus su
30 161.4 13.3 198743 1 BSUB0006 BSUB0006 Bacillus su
31 160.2 13.2 300275 1 AE016751 AE016751 Staphyloc
32 159.6 13.1 302050 1 AL935256 AL935256 Lactobaci
33 159.2 13.1 349814 1 CR378673 CR378673 Photobact
34 156.8 12.9 153950 8 AC104429 AC104429 Oryza sat
35 156.4 12.9 609 6 CO871473 CO871473 Sequence
36 156.4 12.9 1148 8 AY017216 AY017216 Saccharom
37 156.4 12.9 1160 6 AK179484 AK179484 Sequence
38 155.6 12.8 300000 1 AP005073 AP005073 Vibrio pa
39 155.6 12.8 2372 3 AK173431 AK173431 Ciona int
40 155.6 12.8 2378 3 AB036852 AB036852 Ciona int
41 155.6 12.8 2380 3 AK173392 AK173392 Ciona int
42 155.2 12.8 1848 4 AY618932 AY618932 Oryctolag
43 155.2 12.8 147205 8 AC146702 AC146702 Genomic s
44 154.8 12.7 2080 8 AB049836 AB049836 Saccharom
45 153.8 12.6 1021 11 CNS06JUF AL401665 T7 end of

ALIGNMENTS

RESULT 1
BT009518 928 bp mRNA linear PLN 20-JUN-2003
LOCUS Triticum aestivum clone wrl.pk0101.e2:figs, full insert mRNA
DEFINITION sequence.
ACCESSION BT009518
VERSION BT009518.1 GI:32129069
KEYWORDS FLI CDNA.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 928)
AUTHORS Tingey,S.V., Walters,P., Powell,W., Dolan,M., Miao,G.-H.,
Caraher,N.R., Hanafey,M.K. and Hainey,C.F.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2003) Crop Genetics, E. I. Dupont de Nemours and
Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
USA

FEATURES
source

Location/Qualifiers
1..928
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wrl.pk0101.e2:figs"

ORIGIN

Query Match 29.5%; Score 359.6; DB 8; Length 928;
Best Local Similarity 65.5%; Pred. No. 1.6e-76;
Matches 558; Conservative 0; Mismatches 289; Indels 5; Gaps 2;
QY 367 GCCCGGGAGCCCGCAGCCAGTCAAGAGAGCGCTGTAATGTCGAACATTTGGGA 426
DB 77 GCCCGCTAATGGATCAGCCAGGTATCGAAGCTTGTGACCTCAACTGTTGGGAA 136
QY 427 ATCGACTAATATTTTATGGCAATTTGCTTGTGACCAATCTGATAGACAGAAATTTGCT 486
DB 137 ATCGAACAACGTTCTTTGGCATGACTGTCCTAATAGTCACTTTGAGAGCGGAACTGCT 196
QY 487 GGGACAAAAGCGTGTGCTGATGATGAACAGGACTCAGTGTTCAGGGAAAAGTACTCT 546

Db 197 AAATCAGAAGGGTTGTGTGTGGATACACGGGTTAAGTGGTTTCAGGGAAAAAGCACACT 256
QY 547 TGCATGTGCACGTAGTCGTGAGTTCATTCAGAGAGCCACCTCAGTATGATCTTATGCG 606
Db 257 AGCATGCGGCGTAAAGTCGAGAGTGCACCTCCAGAGGTCACTGACCTACATTCCTAGACGG 316
QY 607 TGACAACTTCAGACATGCGCTAAATAGAGATTAAAGCTTTAAGGCGAGAGACCGTGCAGA 666
Db 317 TGACAACTTAAGCATGGGTTAAACCGAGACCTCTGTTTCGAGCAAGGACCGTGCCTGA 376
QY 667 AAATATACGAAGAGTGTGGAAGTGGCAAGCTTTTTCGTGATGCTGTGTCTCATATGAT 726
Db 377 AAATATACGAGAGTAGGAGAGTAGCAAGCTGTTTTCAGATGCTGTGTCTCATATGAT 436
QY 727 TCGTAGCTGATATCCATCAGAGAGATCGTGTGATGATGCGCTGCTCTACATCCACA 786
Db 437 TCGTAGCTGATATCACCTCAGAGAGTGAACGCGCTTGCCGCAATTACTGCACAA 496
QY 787 TTCTAACTTTATGAAGTATTTATTTGATTTGCCCTTAAAAATTTCTGAAGCTCGTGATCC 846
Db 497 TTCTACATTCATCGAGTGTGTTTGAATGCTCCACTTGAAGTTTGTGAAGCTAGGATCC 556
QY 847 TAAAGGCTTATCAAGCTTGCACGTACAGGAAGATTAAAGTTTTCATCGGAATGATGA 906
Db 557 AAAAGGCTTGTACAAGCTTGCCGCTGCAGGAAAAATCAAGGGTTTACTGGAAATGATGA 616
QY 907 TCCATACGACCAACCAATTAATGCTGAGATAGTAAATTAAGATGAAGATGAGGATGCC 966
Db 617 TCCATATGAAGCACCTTCTGATGCGAGATAGTATACAGTGCAGGCTGCTGATCGCG 676
QY 967 TTCACCCAAAGCAATGCGCAAGCTTCTATGCTACCTTGAAGAAAAACGATATTGCA 1026
Db 677 CAGCGCTAATCGATGCTGATCAAGTTGTGTCATATCTTGAAGCAATGATGCTTCTTACA 736
QY 1027 AGCTTGG---TATATGATTTTGAAGAGATGATGATCTGATCTTGTGTGTCATTTCTGT 1083
Db 737 GGAATAGAGACGTATGCTATGATGAAAGAAACATTCGAAATGGATGCGCAAGGATGT 796
QY 1084 GGACACATAAGATCTGTTGTTGTCACATGAATAA--AAGGCATCAACATGTAGGAAGT 1141
Db 797 GAAATATGAGGTAGTATTTATCTCTAGAAAGATGATGATATGATGAGACATATATATT 856
QY 1142 AACAGAAGGTAGCGTTTCATTCAGAAACGGATATGGATTTCATTCGTTTAAAAAAGAAAA 1201
Db 857 GACATAAAGATCGAATCTGTACATCATATATATAATGAAATGTTTTCAGCCAAAAA 916
QY 1202 AAAAAAAGAAAA 1213
Db 917 AAAAAAAGAAAA 928

RESULT 2
AKO71285
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J023086D14, full insert sequence.
DEFINITION AKO71285 linear mRNA PLN 24-JUL-2003
ACCESSION AKO71285
VERSION AKO71285.1 GI:32981308
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,

Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, S., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Ootomo, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 1434)
ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., CARNINCI, P., DOI, K., FUJIMURA, T., FUKUDA, S., HANAGAKI, T., HARA, A., HASHIDUME, W., HAYASHIDA, K., HAYASHIZAKI, Y., HAYATSU, N., HIRAMOTO, K., HIRAOKA, T., HORI, F., HOTTA, I., IIDA, J., IIDA, Y., IKEDA, R., IMAMURA, K., IMOTANI, K., ISHIBIKI, J., ISHII, Y., ISHIKAWA, M., ITOH, M., KAGAWA, S., KATOH, H., KAWAGASHIRA, N., KAWAI, J., KAWAMATA, M., KIKUCHI, S., KISHIKAWA-HIROZANE, I., KISHIMOTO, N., KOBAYASHI, M., KODAMA, T., KOJIMA, K., KOJIMA, Y., KONDO, S., KONNO, H., KOUNDA, M., KOTA, S., KURIHARA, C., KUROSAKI, T., KUSUMEGI, T., LI, C., LU, M., MASUDA, H., MATSUBARA, K., MATSUYAMA, T., MIURA, J., MIYAZAKI, A., MIZUNO, K., MURAKAMI, K., MURATA, M., NAGATA, T., NAKAMURA, M., NAMIKI, T., NARIKAWA, R., NIIKURA, J., NISHI, K., NOMURA, K., NUMASAKI, R., OHNEDA, E., OHNO, M., OHTSUKI, K., OKA, M., OOKA, H., OSATO, Y., OTO, Y., OTOMO, Y., RYU, R., SAITOH, H., SAKAI, C., SAKAI, K., SAKAZUME, N., SANO, H., SASAKI, D., SATO, K., SATOH, K., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T., SHISHIKI, T., SOGABE, Y., SUGANO, S., SUGIYAMA, A., SUZUKI, K., SUZUKI, Y., TAGAMI, M., TAGAMI, T., TAKEDA, Y., TAGAWA, A., TAKAHASHI, F., TAKAKU-AKABIRA, S., TANAKA, T., TOMARU, A., TOYA, T., TSUNODA, Y., UEDA, M., WAKI, K., XIE, Q., YAHAGI, W., YAMADA, H., YAMAMOTO, M., YASUNISHI, A., YAZAKI, J., YOKOMIZO, S. and YOSHIMURA, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashidume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, C., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Soabe, Y., Tagami, M., Tagami, T., Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1..1434
/organism="Oryza sativa (japonica cultivar-group)"

ORIGIN	/mol_type="mrna" /cultivar="Nipponbare" /db_xref="taxon:39947" /clone="J023086D14"									
	Query Match	27.4%;	Score	333.2;	DB 8;	Length	1434;			
	Best Local Similarity	71.6%;	Pred. No.	4.8e-70;						
	Matches	437;	Conservative	0;	Mismatches	173;	Indels	0;	Gaps	0;
QY	424	GAATTCGACATATATTTATGCGACAAATGCTTGATGAGCAATCTGATAGACAGAAATT	483							
Db	462	GAAGCGTCCTCAATATCTCTGCGATGATGTGCGAGTTGGCCAGGCTGATCGGCAGAGCT	521							
QY	484	GCTGGACAAAAGGCTGTCGTATGGATACAGAGCTCAGTGTTCAGGGAAGTAC	543							
Db	522	ACTGAAGCAGAAGGTGGTTGTTGGATCAGAGCTTAGTGGTTAGGTAAGTAC	581							
QY	544	TCTTGATGTGCACTGAGTCGTGAGTTGCATTCGAGAGGCCACCTCAGTATGTAATTGA	603							
Db	582	CTGGCATGCACATTAGATCGAGAGCTCCATACAGAGGGAAGCTTCTTATGTTCTTGA	641							
QY	604	TGGTGACAACTCAGACATGGCTAAATAGAGATTAAAGCTTTAAGGAGAACCGTGC	663							
Db	642	TGGTGATAATTTAAGACATGGTTTGAACAAGGATCTTGCTTTAAGGCGGAAGACCGTGC	701							
QY	664	AGAAATATACGAAGTGGTGAAGTGGCAAGCTTTTGGTGTGATGCTGGTGTGATATG	723							
Db	702	TGAATAATATACGACAGTGGTGGAGTAGCAAAAGCTATTTCGCAGATGCGAGCCTAGTATG	761							
QY	724	CATTGCTAGCTTGATATCTCCATACAGAGAGATCGTGATGATGCCGTGCTCTACTTCC	783							
Db	762	CATTGCAAGTTTCATATCTCCGTATAGAGAGACCGTGAGTCTTGCTGTCATATTGTC	821							
QY	784	ACATTCTAATTTATGAAGTATTTATGATTTGCCCTTAAATAATTTGTGAAGCTCGTGA	843							
Db	822	AGATGGTAGCTTTATTTGAAGTTTCTTGAACATGCCCTTGGAAATTTGTGAGTCAAGGA	881							
QY	844	TCCTAAGGCTTATACAGCTTGACGTACAGGAAGATTAAAGTTTCACTGGAAATGA	903							
Db	882	TCCTAAGGCTTATATAAGCTTGCTGCGTAGGAAAAATAAAGGTTTATCTGGAATTTGA	941							
QY	904	TGATCCATACGAACCAACCAATTAATGTTGAGATAGTAATTAAGATGAAGATGAGGAATG	963							
Db	942	TGACCTTATGATCACCATTAACTCTGAGATTGAGATCAAGAGTGGATGGTGTATG	1001							
QY	964	CCCTTACCCAAAGCAATGGCCAAAGTCTTATGTCCTTGAAGAAAAACGGATATTT	1023							
Db	1002	CCCTTACCATCGGACATGGCAGGACAAGTAGTCACTTATCTAGAGGAGAAAGGGTTTCT	1061							
QY	1024	GCAGCTTAG	1033							
Db	1062	GCATGATTAG	1071							
RESULT 3	AF178976									
LOCUS	AF178976									
DEFINITION	Zea mays adenosine-5'-phosphosulfate kinase (AK1) mRNA, partial cds.									
ACCESSION	AF178976									
VERSION	AF178976.1									
KEYWORDS										
SOURCE	Zea mays									
ORGANISM	Zea mays									
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.									
REFERENCE	1 (bases 1 to 970)									
AUTHORS	Bolchi, A., Petruccio, S. and Ottonello, S.									
TITLE	Isolation and comparative expression analysis of a maize cDNA encoding adenosine 5'-phosphosulfate kinase									
JOURNAL	Unpublished									
REFERENCE	2 (bases 1 to 970)									
AUTHORS	Bolchi, A., Petruccio, S. and Ottonello, S.									
TITLE	Direct Submission									
JOURNAL	Submitted (19-AUG-1999) Istituto di Scienze Biochimiche, Universita' di Parma, Viale delle Scienze, Parma 43100, Italy									
FEATURES	Location/Qualifiers									
source	1..970									
	/organism="Zea mays"									
	/mol_type="mrna"									
	/db_xref="taxon:39947"									
gene	<1..970									
	/gene="AK1"									
CDS	<1..868									
	/gene="AK1"									
	/EC_number="2.7.1.25"									
	/note="APS kinase; ATP adenosine-5'-phosphosulfate 3'-phosphotransferase"									
	/codon_start=2									
	/product="adenosine-5'-phosphosulfate kinase"									
	/protein_id="AAP17236.1"									
	/db_xref="GI:6563285"									
	/translation="VVPVLLLLAVARRAPEPEPRKSRVASAPPPPTSPGMNQ NEGRRLKRGTPVECTGDRSVEQEHAGVEGKAWKMSVTPKSSNIFWHDCPVGK TDRONVLKQGVVMTGLSGSGKSTLACTGLRELHTRGKLVAVLDGNLRLGNKDL GFKAEDRAENIRRVVAKLPADAGLVCIASLISPHRRDRESCEALLSDSFSFLEVLNMS LELCEARDPGKLYKLARAGKITGFTGDDPYEAPLNCIEIEIKVDGVCVPPPAEMAGQV VTYLEEKGFUHE"									
ORIGIN										
	Query Match	25.6%;	Score	311;	DB 8;	Length	970;			
	Best Local Similarity	70.0%;	Pred. No.	1.1e-64;						
	Matches	435;	Conservative	0;	Mismatches	180;	Indels	6;	Gaps	1;
QY	413	TCGAACATTTGGAAATCGACTAATATTTTATGGCACAAATTCCTTGAATGGACAATCTGAT	472							
Db	254	TCCACCGTGGCGAAGTCATCAATATCTTCTGGCACGATTCGCCAGTGGCAAGACTGAT	313							
QY	473	AGACAGAAATTTGCTGGGACAAAAGGCTGTCTGTATGGATAACAGGACTCAGTGGTTCA	532							
Db	314	CGCCAGAACGCTACTCAAGCAGAAAAGGCTGTGTTCTCTGGATTACAGGCTTAGTGGTTCA	373							
QY	533	GGGAAAAGTACTCTTGTGATGTGCATGCTGAGTCGATGCTGAGTTGTCATTCAGAGGCCACCTCAG	592							
Db	374	GGTAAAGTACTCTTGGCATGTACATTAGGCCGTGAGCTCCACAGAGGAGAGCTTCA	433							
QY	593	TATGTACTTGTGATGATCAACCTCAGACATGGCCCTAAATAGAGATTTTAAGCTTTAAGGCA	652							
Db	434	TATGCTCTTGAAGTATTAACCTTAAGACATGCTCTGAACAAGGATCTTGGCTTCAAGCT	493							
QY	653	GAAGACCGTGCAGAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATGCT	712							
Db	494	GAAGACCGTGCAGAAAATATTAAGGAGAGTT-----GTAGCAAAAGTTATTTCAGATCG	547							
QY	713	GGTGTATATGCTAGTTGTAGCTTGCATATCTCCATACAGGAGAGATCGTGATGCTCGCT	772							
Db	548	GGCCTTGTATGTATGCAAGTTTGTATATCTCCATAGGAGAGACCGTGAATCTTGGCT	607							
QY	773	GCTTACTTCCACATTTCTAACTTTTATGAAGTATTTTATGATTTTGGCCCTTAAATAATTTGT	832							
Db	608	GCATTGTTGTGATAGTAGCTTCAATTTGAAGTTTCTTGAACATGTCCTTGAATTAATGT	667							
QY	833	GAAGCTCGTATCTTAAAGGCTTATACAGCTTGCACGTACAGGAAGATTAAGGTTTC	892							
Db	668	GAAGCAAGAGATCCAAAGGGCTCTATAAGCTTCTCTGTCAGGAAAAATAAAGGGTTT	727							
QY	893	ACTGGAATGATGATCCATACGACCACTTAATGCTGAGATAGTAAATTAAGATGAA	952							
Db	728	ACAGGAATAGATGACCTTTATGAGCACCCTGAAATTTGTGAGATTGAGATCAAGAGATA	787							
QY	953	GATGAGGAATGCCCTTCCACCCAAAGCAATGCCCAAGCAAGTTCTATGCTACTCTTGAAGAA	1012							
Db	788	GATGCTGATGCCCTCCGCCCTGCTGAGATGGCAGGCAAGTTGTTACTTACCTTGGAG	847							

```
QY 1013 AACGGATATTTGCAAGCTTAG 1033
|||||
Db 848 AAAGGCTTCTGCACGAGTAG 868

RESULT 4
AF044285 1086 bp mRNA linear PLN 11-AUG-1998
Catharanthus roseus adenosine-5'-phosphosulfate-kinase (Crakn)
mRNA, complete cds.
ACCESSION AF044285
VERSION AF044285.1 GI:2832299
SOURCE Catharanthus roseus (Madagascar periwinkle)
ORGANISM Catharanthus roseus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae;
Vinceae; Catharanthus.
1 (bases 1 to 1086)
Arz. H.E., Gieselmann, G., Schiffmann, S. and Schwenn, J.D.
A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana
Biochim. Biophys. Acta 1218 (3), 447-452 (1994)
94325358
PUBMED 8049272
REFERENCE 2 (bases 1 to 1086)
Schiffmann, S. and Schwenn, J.D.
Isolation of cDNA clones encoding
adenosine-5'-phosphosulfate-kinase (BC2.7.1.25) from Catharanthus
roseus (Accession No. AF044285) and an isoform (akn2) from
Arabidopsis (Accession No. AF043351) (PGR98-116)
Plant Physiol. 117 (3), 1125 (1998)
3 (bases 1 to 1086)
Schiffmann, S. and Schwenn, J.D.
Direct Submission
Submitted (23-JAN-1998) Biochemie der Pflanzen, Ruhr-Universitaet
Bochum, Universitaetsstrasse 150, Bochum 44780, Germany
Location/Qualifiers
1..1086
/organism="Catharanthus roseus"
/mol_type="mRNA"
/db_xref="taxon:4058"
/sub_clones="KSCRakn"
/note="heterotrophic cell culture grown under sulfate
deficiency"
1..1086
/gene="Crakn"
39..977
/gene="Crakn"
/EC_number="2.7.1.25"
/function="phosphorylates 3'-OH group of
adenosine-5'-phosphosulfate"
/note="APS-kinase; putative processing site after residue
46"
/codon_start=1
/product="adenosine-5'-phosphosulfate-kinase"
/protein_id="AAC31145.1"
/db_xref="GI:2832300"
/translation="MIGSVKRPVSVCLPEFDFTSTGLGKSSVKLPVNFAGPGSG
GGEVLGLFAPIKATEGKTSFQVNGKVDQFRHLQPSDCNSDSSLNCGPFGKK
ILQTTTNGSNILHMKCAVEKSERQEPLOQRGCVITWITGLSGKSTLACALSRGLH
AKGKLTLYLDGDNVRHGLNSDLSFADRAENIRIRIGVAKLFADAGVICIASLISPY
RKPPDACRSLPPEGDFIEVFDVPLKVCEARPKLYKLARAGIKGFTGIDDPYEP
LKSEVLHQLGMCDSPCDLADIVISYLEENGYLKA"

ORIGIN
Query Match 25.4%; Score 309; DB 8; Length 1086;
Best Local Similarity 67.4%; Pred. No. 3.5e-64;
Matches 435; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 388 CCACGTGAAGGAGGAGCTGTAAATCGAACAATCGAATAATATTATTTATGGCA 447
|||||
```

```
Db 332 CCCAGTAAAAAGATATTTACAAACAACCTACCGTTGGAAACTCTACTATATACCTTTGGCA 391
QY 448 CAATTGCTTGTATTGGACAAATCTGATAGACAGAAATTCGTGGAGCAAAAAGGCTGTGTGCT 507
|||||
Db 392 TAAATGTGCTGTGGAAAAAGTGAAGGACAGAAACCCCTGCAGCAAAAGGGGCTGTGTAT 451
QY 508 ATGGATAACAGGACTCAGTGGTTTCAGGGAAGAAGTACTCTTTCATGTGCACCTGAGTCGTGA 567
|||||
Db 452 ATGGATTACTGTCTTAGTGGTTTCAGGAAAGAGACACCTTGGCATGTGCTTTGAGTCGAGG 511
QY 568 GTTGCAATTGACAGAGCCACCTCAGTATGTACTTGTATGGTGTGACAACTTCACACATGGCCT 627
|||||
Db 512 CTTACATGCAAGGGAACCTAACCTACTCTTTGTATGGTGTATATGTTCACACATGGATT 571
QY 628 AAATAGAGATTTAAGCTTTAAGGCAGAGACCGTGCAGAAATATACGAACAGTTGGTGA 687
|||||
Db 572 GAACAGTGAUCTTTAGTTTTAAAGCAGAAAGATCGAGCGGAGAAATATTAGGAGATTGGAGA 631
QY 688 AGTGGCAAAAGCTTTTGTGATGCTGGTGTGCATATGCATTTGCTAGCTTTGATATCTCCATA 747
|||||
Db 632 GTTTGAAGGCTCTTTGCTGATGCTGGAGTCAATTTGCATTCGAAAGTTTGATATCTCCCTA 691
QY 748 CAGGAGAGATCGTATGATGATGCCGTGCTTACTTCCACATTTCTAACTTTATTGAAGTATT 807
|||||
Db 692 CAGAAAGCCCCCGATGCTGCTCTCTTTGCTTCCAGAAAGGGGATTTCAATTGAGGTATT 751
QY 808 TATTGATTGCCCCCTTAAATTTCTGAAGCTCGTGCATCTTAAAGSCCTATACAAAGCTTGC 867
|||||
Db 752 CATGGATGTCCTCTGAAAGTCTCGAGGCAAGGAGCCCTAAGGGCTTATATAAGCTTGC 811
QY 868 ACGTACAGGAAAGATTAAAGGTTTTCACTGGAAATTTGATGATCCATACGAAACCAATTTAA 927
|||||
Db 812 ACGAGCTGGTAAATCAAAGGCTTTACTGGAATTTGATGATCCTTACGAGCCGCCACTTAA 871
QY 928 TGGTGAATAGTAAATTAAGATGAAGATGAGGAATGCCCTTACCCAAAGCAATGGCCAA 987
|||||
Db 872 ATCAGAGATTGTTTACATCAGAAAGCTTGGTATGTGTGATTTCTCCATGTGATTTGGCTGA 931
QY 988 GCAAGTTCTATGCTACCTTGAAGAAACCGGATATTTTGAAGCTTA 1032
|||||
Db 932 TATAGTGATATCATATCTGGAGGAATGATATCTTAAAGGCATA 976

RESULT 5
AV085264 1068 bp mRNA linear PLN 14-APR-2003
LOCUS Arabidopsis thaliana clone 14216 mRNA, complete sequence.
DEFINITION Arabidopsis thaliana clone 14216 mRNA, complete sequence.
ACCESSION AV085264
VERSION AV085264.1 GI:21403974
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1068)
Haas, B.J., Volkovsky, N., Town, C.D., Troukhan, M., Alexandrov, N.,
Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. 3 (6), RESEARCH0029 (2002)
22088475
PUBMED 12093376
REFERENCE 2 (bases 1 to 1068)
Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
3 (bases 1 to 1068)
Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
Feldmann, K.
Direct Submission
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
```

Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the ws or laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genet carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

source

1. .1068
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="14216"
94. .924
/codon_start=1
/product="putative adenosine phosphosulfate kinase"
/protein_id="AA062496.1"
/db_xref="GI:21553403"

CDS

24.4%; Score 296.4; DB 8; Length 1068;
Best Local Similarity 68.0%; Pred. No. 4.1e-61;
Matches 430; Conservative 0; Mismatches 196; Indels 6; Gaps 1;
411 TGTGAAACATTGGGAAATCGACTATATATTTATGGCACAATTCCTTGTATGGACAATCTG 470
308 TGTCTACGGTCGGAAACTCGACAAATATAAAGTGGCATGAATGTCTCTGTGAGAAAGTTG 367
471 ATAGACAGAAATTCGTGGACAAAAGCGTGTGCTATGGATACAGCACTCAGTGGTT 530
368 ATAGACAGAGATTCGTTGATCAGAAGGATGTGATTTGGGTACCGGTCTTAGTGGTT 427
531 CAGGGAAGAGTACTCTTGATGTGCATCTGCACTGAGTTCGTGATTCGAGAGCCACCTCA 590
428 CAGGGAAGAGTACTTTGGCTTGTGCTTTGAATCAGATGTTGATCAAAAGGGAAGCTTT 487
591 COTATGTACTTGAATGGTGACACCTCAGACATGGCTTAATAGAGATTAAGCTTTAAGG 650
488 GTTATATTTCTGATGTGATAATGTTAGGCATGGCTTAACCGTGCATCTTAGCTTTAAAG 547
651 CAGAGACCGTCGACAAATATACGAGAGTTGGTGAAGTGGCAAGCTTTTGTGCTGATG 710
548 CTGAGGATCGTCAGAGAAATATTCGTAGAGTTGGAGAGTGTCTAAGCTTTTTCGGATG 607
711 CTGGTGTTCATATGCAATTCGTAGCTATCTCCATACAGAGAGATCGTGATGATGCC 770
608 CTGGAATATCTGCATTCGGAGTTGATATCTCTTATAGACATAGGACGCTTGTG 667
771 GTGCTCTACTTCCACATTCCTAATTTATGAAGTATTTATGATTTGCCCTCAAAAATTT 830
668 GAAGTTTGTCTCCCGAGGAGATTTTGTAGGTGTTGATGATGATGATGATGATGATG 727
831 GTGAGCTCTGTATCTTAAGGCTTATACAGCTTGCACGTACAGGAAGATTAAGCTT 890
728 GCGAGCGGAGGATCAAAAGGGTCTTTCAAGCTTGCCTGTCGACGGAAGATCAAGGTT 787
891 TCACTGGAATTGATGATCCATACGAACCAATTAATGATGATGATGATTAAGATGA 950

ORIGIN

Query Match 24.4%; Score 296.4; DB 8; Length 1068;
Best Local Similarity 68.0%; Pred. No. 4.1e-61;
Matches 430; Conservative 0; Mismatches 196; Indels 6; Gaps 1;
411 TGTGAAACATTGGGAAATCGACTATATATTTATGGCACAATTCCTTGTATGGACAATCTG 470
308 TGTCTACGGTCGGAAACTCGACAAATATAAAGTGGCATGAATGTCTCTGTGAGAAAGTTG 367
471 ATAGACAGAAATTCGTGGACAAAAGCGTGTGCTATGGATACAGCACTCAGTGGTT 530
368 ATAGACAGAGATTCGTTGATCAGAAGGATGTGATTTGGGTACCGGTCTTAGTGGTT 427
531 CAGGGAAGAGTACTCTTGATGTGCATCTGCACTGAGTTCGTGATTCGAGAGCCACCTCA 590
428 CAGGGAAGAGTACTTTGGCTTGTGCTTTGAATCAGATGTTGATCAAAAGGGAAGCTTT 487
591 COTATGTACTTGAATGGTGACACCTCAGACATGGCTTAATAGAGATTAAGCTTTAAGG 650
488 GTTATATTTCTGATGTGATAATGTTAGGCATGGCTTAACCGTGCATCTTAGCTTTAAAG 547
651 CAGAGACCGTCGACAAATATACGAGAGTTGGTGAAGTGGCAAGCTTTTGTGCTGATG 710
548 CTGAGGATCGTCAGAGAAATATTCGTAGAGTTGGAGAGTGTCTAAGCTTTTTCGGATG 607
711 CTGGTGTTCATATGCAATTCGTAGCTATCTCCATACAGAGAGATCGTGATGATGCC 770
608 CTGGAATATCTGCATTCGGAGTTGATATCTCTTATAGACATAGGACGCTTGTG 667
771 GTGCTCTACTTCCACATTCCTAATTTATGAAGTATTTATGATTTGCCCTCAAAAATTT 830
668 GAAGTTTGTCTCCCGAGGAGATTTTGTAGGTGTTGATGATGATGATGATGATGATGATG 727
831 GTGAGCTCTGTATCTTAAGGCTTATACAGCTTGCACGTACAGGAAGATTAAGCTT 890
728 GCGAGCGGAGGATCAAAAGGGTCTTTCAAGCTTGCCTGTCGACGGAAGATCAAGGTT 787
891 TCACTGGAATTGATGATCCATACGAACCAATTAATGATGATGATGATTAAGATGA 950

788 TTACCGGATCGATGACACCTTACGAGCCACCATGAACTGCGAG-----ATTCTCTAG 841
951 AAGATGAGGATGCCCTTACCCAAAGCAATGGCCCAAGATTTCTATGCTACCTTGAAG 1010
842 GACGTGAAGGAGGAACTTCTCTATCGAATGGCGGAAAGGTCGTCGATATCTAGATA 901
1011 AAAACGGATATTTGCAAGCTTAGTATATATAT 1042
902 ACAGGGTTATCTTCAAGCATACATATCTTCT 933
ATU05238 1077 bp mRNA linear PLN 23-JUN-1998
Arabidopsis thaliana APS kinase mRNA, complete cds.
U05238
VERSION U05238.1 GI:450234
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1077)
AUTHORS Jain, A. and Leustek, T.
TITLE A cDNA clone for 5'-adenylylphosphosulfate kinase from Arabidopsis thaliana
JOURNAL Plant Physiol. 105 (2), 771-772 (1994)
MEDLINE 94345022
PUBMED 8066145
REFERENCE 2 (bases 1 to 1077)
AUTHORS Lee, S. and Leustek, T.
TITLE APS kinase from Arabidopsis thaliana: genomic organization, expression, and kinetic analysis of the recombinant enzyme
JOURNAL Biochem. Biophys. Res. Commun. 247 (1), 171-175 (1998)
MEDLINE 98300303
PUBMED 9636674
REFERENCE 3 (bases 1 to 1077)
AUTHORS Leustek, T.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-1994) Thomas Leustek, Center for Agricultural Molec. Biology, Rutgers University, Cook College, College Farm Road, New Brunswick, NJ 08903, USA
FEATURES
source Location/Qualifiers
1. .1077
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="ATK"
/tissue_type="whole plant"
33. .863
/EC_number="2.7.1.25"
/function="phosphorylates 5'-adenylylsulfate on 3'-hydroxyl group"
/note="similar to yeast APS kinase (MET14); SwissProt Accession Number Q02196; similar to Escherichia coli APS kinase (CysC): SwissProt Accession Number P23846; adenosine 5'-phosphosulfate kinase"
/codon_start=1
/evidence="experimental"
/product="APS kinase"
/protein_id="AAC50035.1"
/db_xref="GI:450235"
transit_peptide 33. .143
/note="potential chloroplast transit peptide with a protease cleavage site"

2

```
QY 531 CAGGGAAGAGTACTCTTGCAATGTGCACTGAGTCGTGAGTTGCAATGCGAGAGCCACCTCA 590
Db 427 CAGGGAAGAGTACTCTTGCTTGGCTTGGCTTGAATCAGATGTTGATCAAAAGGGGAAGCTTT 486
QY 591 CTTATGCTACTGTCATGTCACACCTCAGACATGGCTAAATAGAGATTTAGCTTTAAGG 650
Db 487 GTTATATCTTGATGGTGAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAAG 546
QY 651 CAGAAGACCGTCGAGAAAATATACGAAGAGTTGGTCAAGTGGCAAGCTTTTGTGCTGATG 710
Db 547 CTGAGGATCGTCAGAGATATTCGTAGAGTTGGAGAGTTGCTAAGCTTTTGGCGATG 606
QY 711 CTGCTGTCATATGCAATGCTAGCTGATATCTCCATACAGAGAGATCGTGATGATGCC 770
Db 607 CTGGAATAATCTGCATTTGCGAGTTTGATATCTCTTATAGAACAGATAGGAGCGCTTGTG 666
QY 771 GTGCTCTACTCCCATCTTAATCTTATTAAGATTTATTTGATTTGGCTTGGCTTAAATTT 830
Db 667 GAAGTTTGTCTCCCGAGGAGATTTTGTGAGGTGTTGATGATGATCGCTTAGTGT 726
QY 831 GTGAAGCTCGTGATCTTAAAGCGCTATACAAAGCTTTCACAGCTTGCAGGAAAGATTAAAGTT 890
Db 727 GCGAGCGAGGATCCAAAGGCTCTTACAGCTTGTCTGTCGACGAAAGATCAAGGTT 786
QY 891 TCATCGGAATTGATGATCATACGACCAACCAATTAATGTTGAGATAGTAATTAAGATGA 950
Db 787 TTACCGGATCGATGACCTTACGAGCGACCATTTGAATCGCGAG-----ATTCTCTAG 840
QY 951 AGATGAGGAATGCCCTTACCACCAAGCAATGCCCAAGCAAGTTCTATGCTACTCTTGAAG 1010
Db 841 GACGTGAAGGAGGAATCTCTCTATCGAAATGGCGGAAAGTCTCGGATACTTTAGATA 900
QY 1011 AAAACGGATATTTCGAAGCTTAGTATATGAT 1042
Db 901 ACAAGGTTATCTTCAAGCATTAACATACTTCT 932
```

RESULT 8

```
ATCLAPSK
LOCUS ATCLAPSK 1185 bp mRNA linear PLN 05-FEB-1998
DEFINITION A.thaliana (L.Heynh.) chloroplast mRNA for recombinant APS-kinase.
ACCESSION X75782
VERSION X75782.1 GI:414736
KEYWORDS APS-kinase; ATP:adenylylsulfate-3'-phosphotransferase.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE 1
AUTHORS Arz,H.E., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.
TITLE A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana
JOURNAL Biochim. Biophys. Acta 1218 (3), 447-452 (1994)
MEDLINE 94325358
PUBMED 8049272
REMARK (sites)
2 (bases 1 to 1185)
AUTHORS Arz,H.E., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.
TITLE A chloroplast APS-kinase cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1185)
AUTHORS Schwenn,J.D.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-1993) Schwenn J. D., Ruhr University Bochum,
Biology, Universitaetsstr. 150, 44780 Bochum, Germany
FEATURES
Location/Qualifiers
1..1185
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="FM239X14cDNA"
/tissue_type="leaf"
/clone_lib="cDNA in lambda ZapII"
```

```
CDs
/dev_stage="vegetative"
42..872
/EC_number="2.7.1.25"
/function="phosphorylates 3'-OH group of adenylylsulfate"
/standard_name="ATP:adenylylsulfate-3'-phosphotransferase"
/citation=[2]
/codon_start=1
/product="APS kinase"
/protein_id="CAA53426.1"
/db_xref="GI:414737"
/db_xref="COA:Q43295"
/db_xref="UniProt/Swiss-Prot:Q43295"
/translation="MIAAGAKSLGLSMASPKGIFDMSMSNSRSVVVRAVCVMDGSG
QTLSHNKGSGTPEVKSNGHTGQKQGPLSTVGNSTNFKWHECSVEKVDQRLLDQKGC
VIWYTLGSGSKSLACALNOMLYQKGLCYLILDNDVNRHGLNDRDLSPKADRAENIR
RVGVAKLFPADAGIICIASLISPYRTDRDACSLLPEGDFVEFMDVPLSVCEARDPK
GLYKLARAGKIKGFTGIDDDPYEPPLNCEISLGRSGGTSPIEMAEKVGWYLDNKGVLQA
"
162..869
/mat_peptide
/product="APS-kinase"
/EC_number="2.7.1.25"
/notes="putative mature peptide only - gene product
starting with ATG at position 162 is enzymatically active,
but maturation in the chloroplast has not yet been shown.
Homology to sequences from E. coli & ye"
/citation=[1]
/citation=[2]
/transit_peptide 42..153
/notes="putative transit peptide with a processing site
suitable for stromal protease; ref: Gavel & Heijne FEBS
Lett. 261, 455, 1990"
984..1012
stem_loop
polyA_site
985..990
stem_loop
1092..1120
ORIGIN
Query Match 24.4%; Score 296.4; DB 8; Length 1185;
Best Local Similarity 68.0%; Pred. No. 4.2e-61;
Matches 430; Conservative 0; Mismatches 196; Indels 6; Gaps 1;
QY 411 TGTGAACATTTGGAAATCGACTAATATTTATGGCACAATGCTTGTGATGGACAACTCG 470
Db 256 TGTCTACGTCGGAACCTCGACAAATATAAAGTGGCATGAATGTTCTGTTGAGAAAGTTG 315
QY 471 ATAGACAGAAATTCCTGGGACAAAAGCTGTGCTGATGATAACAGCACTCAGTGTGT 530
Db 316 ATAGACAGAGATTCCTGATCAGAAAGATGTGATTTGGGTCAACCGTCTTAGTGTGT 375
QY 531 CAGGGAAGTACTCTTTGCAATGTCACGTGAGTCGTGAGTTGCAATGCGAGAGCCACCTCA 590
Db 376 CAGGGAAGTACTCTTTGGCTTGTGCTTTGAATCAGATGTTGATCAAAAGGGGAAGCTTT 435
QY 591 CGTATGTAATGATGGTGAACAACCTCAGACATGGCTTAATAGAGATTTAAGCTTTAAGG 650
Db 436 GTTATATCTTTGATGGTGAATAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAG 495
QY 651 CAGAAGACCGTCGAGAAAATATACGAAGAGTTGGTGAAGTCGCAAGCTTTTGTGCTGATG 710
Db 496 CTGAGATCGTCAGAGATTAATTCGTAGATTTGAGAGGTTGCTAAGCTTTTGGCGATG 555
QY 711 CTGCTGTCATATGCAATGCTAGCTTGTATCTCCATACAGAGAGATCGTGATGATGCC 770
Db 556 CTGGAATAATCTGCATTTGGAGTTTGATATCTCTTATAGAACAGATAGGAGCGCTTGTG 615
QY 771 GTGCTCTACTTCCACATCTTAATCTTATTAAGATTTATTTGATTTGGCTTAAATTT 830
Db 616 GAAGTTTGTCTCCCGAGGAGATTTTGTGAGGTGTTGATGATGATGATCGCTTAGTGT 675
QY 831 GTGAAGCTCGTGATCTTAAAGCGCTTATACAGAGCTTGCACGTACAGGAAAGATTAAAGTT 890
Db 676 GCGAGGCGAGGGATCCAAAGGGTCTTTTACAAGCTTGTCTGTCGAGGAAAGATCAAGGTT 735
QY 891 TCACTGGAATTGATGATCCATACGAAACCAATTAATGTTGATAGTAAATTAAGATGA 950
```

```
Db      736 TTACCGGATCGATGACCCCTTACGAGCCACCAATTGAATGCGAG-----ATTTCTCTAG 789
Qy      951 AAGATGAGGAATGCCCTTACCCAAAGCAATGGCCAAAGATTCTATGCTACCTTGAAG 1010
Db      790 GACGTGAAGGAGAACTTCTCTATCGAAATGGCGAAAGGTCGTCGGATCTTAGATA 849
Qy      1011 AAAACGGATATTTGCAAGCTTAGTATGTAT 1042
Db      850 ACAAGGTTATCTTCAAGCAATACATACTTCT 881

RESULT 9
CQ805858
LOCUS      CQ805858      831 bp      DNA      linear      PAT 10-MAY-2004
DEFINITION      Sequence 2269 from Patent WO2004035798.
ACCESSION      CQ805858
VERSION      CQ805858.1      GI:47111563
KEYWORDS
SOURCE
ORGANISM      Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS      Inze,D.; de Veylder,L. and Vlieghe,K.
TITLE      Identification of novel e2f target genes and use thereof
JOURNAL      CropDesign N.V. (BE)
FEATURES
source      1..831
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match      24.2%; Score 294.4; DB 6; Length 831;
Best Local Similarity      68.3%; Pred. No. 1.2e-60;
Matches 425; Conservative      0; Mismatches 191; Indels      6; Gaps      1;

Qy      411 TGTCAACATTTGGGAAATCGACTAATATTTATGGCACAAATGCTTGATGGACAATCTG 470
Db      215 TGCTACCGTCGGAAACTCGACAAATATAAAGTGGCATGAATGTTCTCTTGAGAAAGTTG 274
Qy      471 ATAGACAGAAATTTGCTGGACAAAAGGCTGTGCTATGGATAACAGCACTCAGTGGTT 530
Db      275 ATAGACAGAGTTGCTTGATCAGAAGAGATGTGTGATTTGGTCACCGGTCTTAGTGGTT 334
Qy      531 CAGGAAAAGTACTCTTGCATGTGCATCGAGTCGTGAGTTGCATTTGCAGAGCCACCTCA 590
Db      335 CAGGAAAGAGTACTTTGGCTTGTGCTTTGAAATCAGATGTTGATCAAAAGGGGAAGCTTT 394
Qy      591 CGTATGTACTTGGTGCACACCTCAGACATGGCTTAATAGAGATTTAAGCTTTAAGG 650
Db      395 GTATATCTTGTATGGTGAATATGTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAAG 454
Qy      651 CAGAAGACCGTCGAGAAAATATACGAAGATTGGTGAAGTGGCAAGCTTTTGTCTGATG 710
Db      455 CTGAGATCGTCGAGAGATATTCGTAGATTGGAGAGTTGCTAAGCTTTTGGCGATG 514
Qy      711 CTGGTGTCAATATGCAATTCAGTATGATATCTCCATACAGAGAGATCGTATGATGCC 770
Db      515 CTGGAATAATCTGCATTCGGAGTTTGATATCTCTTATAGAACAGATAGGACGCTTGTG 574
Qy      771 GTGCTCTACTTCCCAATCTTAACCTTATTAAGTATTTTATGATTTGCCCTTAAAAATTT 830
Db      575 GAAGTTTGTCTCCCGAGGAGATTTTGTGAGGTGTTTCATGGATGTACCGCTTAGTGTGT 634
Qy      831 GTGAAGCTCGTGATCCTTAAAGCCCTATACAAGCTTGACAGTACAGGAAGATTTAAAGTT 890
Db      635 GCGAGCGAGGGATCCAAAGGCTCTTACAGCTTGCTCGTCGCGAGGAAGATCAAAGTT 694
Qy      891 TCATGGAAATTGATGATCCATACGAACCAACCAATTAATGGTGAGATAGTAATTAAGATGA 950
```

```
Db      695 TTACCGGATCGATGACCCCTTACGAGCCACCAATTGAATGCGAG-----ATTTCTCTAG 748
Qy      951 AAGATGAGGAATGCCCTTACCCAAAGCAATGGCCAAAGATTCTATGCTACCTTGAAG 1010
Db      749 GACGTGAAGGAGAACTTCTCTATCGAAATGGCGAAAGGTCGTCGGATCTTAGATA 808
Qy      1011 AAAACGGATATTTGCAAGCTTA 1032
Db      809 ACAAGGTTATCTTCAAGCATA 830

RESULT 10
AX412633
LOCUS      AX412633      831 bp      DNA      linear      PAT 14-JUN-2002
DEFINITION      Sequence 397 from Patent WO0222675.
ACCESSION      AX412633
VERSION      AX412633.1      GI:21445091
KEYWORDS
SOURCE
ORGANISM      Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS      Glazebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.
TITLE      Plant genes, the expression of which are altered by pathogen
infection
JOURNAL      Patent: WO 0222675-A 397 21-MAR-2002;
Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT
CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl,
Jeffrey L. (US) ; Eulgem, Thomas (US)
FEATURES
source      1..831
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match      24.2%; Score 294.4; DB 6; Length 831;
Best Local Similarity      68.3%; Pred. No. 1.2e-60;
Matches 425; Conservative      0; Mismatches 191; Indels      6; Gaps      1;

Qy      411 TGTCAACATTTGGGAAATCGACTAATATTTATGGCACAAATGCTTGATGGACAATCTG 470
Db      215 TGCTACCGTCGGAAACTCGACAAATATAAAGTGGCATGAATGTTCTCTTGAGAAAGTTG 274
Qy      471 ATAGACAGAAATTTGCTGGGACAAAAGGCTGTGCTATGGATAACAGGACTCAGTGGTT 530
Db      275 ATAGACAGAGATTTGCTTGATCAGAAGAGATGTGTGATTTGGGTCACCGGTCTTAGTGGTT 334
Qy      531 CAGGAAAAGTACTCTTGCATGTGCATCGAGTCGTGAGTTGCATTCAGAGGCCACCTCA 590
Db      335 CAGGAAAGAGTACTTTGGCTTGTGCTTTGAAATCAGATGTTGATCAAAAGGGGAAGCTTT 394
Qy      591 CGTATGTACTTGGTGTGACAACCTCAGACATGGCTTAAATAGAGATTTAAGCTTTAAGG 650
Db      395 GTTATATCTTGTATGGTGAATAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAAG 454
Qy      651 CAGAAGACCGTCGAGAAAATATACGAAGATTGGTGAAGTGGCAAGCTTTTGTCTGATG 710
Db      455 CTGAGGATCGTCGAGAGATATTCGTAGATTGGAGAGTTGCTAAGCTTTTGGCGATG 514
Qy      711 CTGGTGTCAATATGCAATTCAGTATGATATCTCCATACAGAGAGATCGTATGATGCC 770
Db      515 CTGGAATAATCTGCATTCGGAGTTTGATATCTCTTATAGAACAGATAGGACGCTTGTG 574
Qy      771 GTGCTCTACTTCCCAATCTTAACCTTATTAAGTATTTTATGATTTGCCCTTAAAAATTT 830
Db      575 GAAGTTTGTCTCCCGAGGAGATTTTGTGAGGTGTTTCATGGATGTACCGCTTAGTGTGT 634
Qy      831 GTGAAGCTCGTGATCCTTAAAGCCCTATACAAGCTTGACAGTTCACGTAAGGAAGATTTAAAGTT 890
```

Db 635 GCGAGCGAGGATCCAAAGGGTCTTTTACAAGCTTGCTCGTGCGAGAAAGATCAAAGGTT 694
 QY 891 TCACCTGGAATTGATGATCCATACGAACCAACCAATTAATGCTGAGATAGTAATTAAGATGA 950
 Db 695 TTACCGGGATCGATGACCTTTAGGACCAACCTTGAACCTGCAG-----ATTCTCTAG 748
 QY 951 AGATGAGGAATGCCCTTCAACCAAGCAATGCCCAAGCAAGTTCTATGCTACCTTTGAAG 1010
 Db 749 GACGTGAAGGAGAACTTCTCTATCGAAATGCGGAAAGGTCGTCGGATACTTAGATA 808
 QY 1011 AAAACGGATATTGCAAGCTTA 1032
 Db 809 ACAAGGGTTATCTTCAAGCATA 830
 RESULT 11
 AX506015
 LOCUS AX506015 831 bp DNA linear PAT 27-SEP-2002
 DEFINITION Sequence 710 from Patent WO0216655.
 ACCESSION AX506015
 VERSION AX506015.1 GI:23387252
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1
 REFERENCE Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
 AUTHORS Stress-regulated genes of plants, transgenic plants containing
 TITLE same, and methods of use
 JOURNAL Patent: WO 0216655-A 710 28-FEB-2002;
 The Scripps Research Institute (US); Syngenta Participations AG
 (CH)
 FEATURES
 source Location/Qualifiers
 1..831
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"
 ORIGIN
 Query Match 24.2%; Score 294.4; DB 6; Length 831;
 Best Local Similarity 68.3%; Pred. No. 1.2e-60;
 Matches 425; Conservative 0; Mismatches 191; Indels 6; Gaps 1;
 QY 411 TCTCGAATCGGAAATCGACTAATATTTATGCAATTCCTTGATGACATCTG 470
 Db 215 TGTCTACGTCGGAACCTCGACAAATATAAAGTGGCATGAATGTTCTGTGAGAAAGTTG 274
 QY 471 ATAGACAGAAATTCGTGGACAAAAGGCTGTGTATGGATACAGGACTCAGTGGTT 530
 Db 275 ATAGACAGAGATGCTTGATCAGAAAGATGTGTATTTGGTACCGGCTTAGTGGTT 334
 QY 531 CAGGGAAGATGCTCTTGATGTCACCTGAGTTCGATTCGAGAGCCACCTCA 590
 Db 335 CAGGGAAGATGCTCTTGCTTTGCTTTGAATCAGATGTTGATCAAAAGGGGAGCTTT 394
 QY 591 CGTATGCTATGATGCTGACACCTCAGACATGGCCCTAATAGAGATTAACCTTTAAGG 650
 Db 395 GTTATATCTTGTATGCTGATAATGTTAGGCATGGCTTAAACCGTATCTTAGCTTTAAG 454
 QY 651 CAGAGACCGTCGAGAAATATACGAAGATGTTGGTGAAGTGCAGAAAGCTTTTGTCTGATG 710
 Db 455 CTGAGGATCGTCAGAGAAATATTCGATAGATGTTGGAGAGTTCCTAAGCTTTTTCGGATG 514
 QY 711 CTGGTGTCTATGCAATTCGATGCTGATATCTCCATACAGGAGATCGTGTGATGCC 770
 Db 515 CTGGAATAATCTGCATTGCGAGTTGATATCTCTTATAGACAGATAGGACGCTTGTG 574
 QY 771 GTGCTCTACTTCACATTTCTAATTTATGAGATATTTATGATTTGCCCTTAAATTT 830
 Db 575 GAAGTTTGTCCCGGAGGAGATTTTGTGAGGTGTTTCATGATGTACCGCTTAGTGGTT 634

QY 831 GTGAAGCTCGTGAATCCTAAAGCCCTATACAAGCTTTCACGTACAGGAAGATTAAGGTT 890
 Db 635 GCGAGCGAGGATCCAAAGGGTCTTTTACAAGCTTGCTCGTGCGAGAAAGATCAAAGGTT 694
 QY 891 TCACCTGGAATTGATGATCCATACGAACCAACCAATTAATGCTGAGATAGTAATTAAGATGA 950
 Db 695 TTACCGGGATCGATGACCTTTAGGACCAACCTTGAACCTGCAG-----ATTCTCTAG 748
 QY 951 AGATGAGGAATGCCCTTCAACCAAGCAATGCCCAAGCAAGTTCTATGCTACCTTTGAAG 1010
 Db 749 GACGTGAAGGAGAACTTCTCTATCGAAATGCGGAAAGGTCGTCGGATACTTAGATA 808
 QY 1011 AAAACGGATATTGCAAGCTTA 1032
 Db 809 ACAAGGGTTATCTTCAAGCATA 830
 RESULT 12
 AX132010
 LOCUS Arabidopsis thaliana At2g14750/F26C24.11 mRNA, complete cds.
 DEFINITION
 ACCESSION AX132010
 VERSION AX132010.1 GI:22135772
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 831)
 REFERENCE Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Banh, J., Bowser, L.,
 AUTHORS Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
 Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
 Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K.,
 Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A.,
 and Ecker, J.R.
 TITLE Arabidopsis ORF clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 831)
 AUTHORS Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Banh, J., Bowser, L.,
 Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
 Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
 Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K.,
 Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A.,
 and Ecker, J.R.
 TITLE Direct Submission
 JOURNAL Submitted (26-JUN-2002) Salk Institute Genomic Analysis Laboratory
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological
 Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA
 COMMENT
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.
 The Salk, Stanford, PGEC (SSP) Consortium members constructed and
 sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim, C.J.,
 Chen, H., Cheuk, R., Shinn, P., Banh, J., Bowser, L., Chang, E.,
 Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B.,
 Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
 Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,
 Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A.,
 and Ecker, J.R.
 Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
 contributed equally to this work as PIs.
 Location/Qualifiers

FEATURES

source	1. .831	LOCUS	BT005193	964 bp	mRNA	linear	PLN 04-MAR-2003
	/organism="Arabidopsis thaliana"	DEFINITION	Arabidopsis thaliana clone U20575 putative adenylylsulfate kinase (At5g67520) mRNA, complete cds.				
CDS	/mol_type="mRNA"	ACCESSION	BT005193				
	/db_xref="taxon:3702"	VERSION	BT005193.1	GI:28827763			
	/chromosome="2"	KEYWORDS	FLI CDNA.				
	/clone="U13881"	SOURCE	Arabidopsis thaliana (thale cress)				
	/ecotype="Columbia"	ORGANISM	Arabidopsis thaliana				
	/note="This clone is in pUNI 51"		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.				
	1. .831	REFERENCE	1 (bases 1 to 964)				
	/note="putative adenosine phosphosulfate kinase"	AUTHORS	Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.				
ORIGIN		TITLE	Arabidopsis Open Reading Frame (ORF) Clones				
		JOURNAL	Unpublished				
	Query Match	REFERENCE	2 (bases 1 to 964)				
	Best Local Similarity	AUTHORS	Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.				
	Matches 425; Conservative 0; Mismatches 191; Indels 6; Gaps 1;	TITLE	Submitted (04-MAR-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA				
	411 TGTCGAACATTCGGAAATCGACTAATATTTATGTCACAAATTCCTTGATTCGACAACTCG 470	JOURNAL	The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.				
	215 TGTCACGGTCGGAACCTCGCAATATTAAGTGGCATGAATGTTCTGTGAGAAAGTTG 274	COMMENT	The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Wong,C., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.				
	471 ATAGACAGAAATTCGTGGACAAAAGCTGTGCTATGGATAACAGGACTCAGTGGTT 530						
	275 ATAGACAGAGATTGCTTGATCAGAAGGATGTGTGATTGGTCCACCGTCTTAGTGGTT 334						
	531 CAGGAAAAGTACTCTTCATGTGCATGACGTGCTGAGTTCGATTCGACAGGCCACCTCA 590						
	335 CAGGAAAGAGTACTTTGGCTTGTGCTTGAATCAGATGTTGTATCAAAAGGGGAAGCTTT 394						
	591 COTATGTACTTCATCGTCACACCTCAGACATGGCTTAAATAGAGATTAAAGCTTTAAGG 650						
	395 GTTATATCTTGATGGTGATATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAG 454						
	651 CAGAACCGTCGACAAAATATACGAAGAGTTGGTGAAGTGCCAAAGCTTTTTCGTGATG 710						
	455 CTGAGGATCGTCGACAGAATATTCGTAGAGTTGGAGAGGTTGCTTAAGCTTTTTCGGATG 514						
	711 CTGGTGTCATATGCTAGCTTGATATCTCCATACAGGAGATCGTGATGCATGCC 770						
	515 CTGGAATAATCTGCATTCGGAGTTTGTGATATCTCTTATAGAACAGATAGGACGCTTGTC 574						
	771 GTGCTCTACTCCACATCTCACTTTATTGAAGTATTTATTGATTTGCCCTTAAATAATTT 830						
	575 GAAGTTGCTCCCGAGGAGATTTTGTGAGTGTTTCATGAGTACCGCTTAGTGTTT 634						
	831 GTGAAGCTCGTGATCCTTAAGCCCTATACAAGCTTTCACGTCACAGGAAGATTAAAGTT 890						
	635 GCGAGCGAGGGATCCAAAGGGTCTTTACAAGCTTTCGTCGACAGGAAGATCAAAGTT 694						
	891 TCACCTGGAATTGATGATCCATACGAACACCAATTAATGGTGAGATAGTAATTAAAGTGA 950						
	695 TTACCGGATCGATGACCTTACGAGGCCACCAATGAACCTGGAG-----ATTTCCTAG 748						
	951 AAGATGAGGAATGCCCTTCCACCAAGCAATGGCCAAAGCAAGTTCCTAGCTACCTTGAAG 1010						
	749 GACGTGAAGGAGAACTTCTCTATCGAAATGCGGAAAGGTCTCGGACTACTAGATA 808						
	1011 AAAACGGATATTGCAAGCTTA 1032						
	809 ACAAGGGTTATCTTCAAGCATA 830						
RESULT 13							
BT005193							

Qy	896	GGAAATTGATGATCCATACGAACCAACCAATTAATGTTGAGATAGTAATTAAGATGAAAGAT	955
Db	824	GGAGTAGATGATCCATATGAAGGCCCTTTGGATTGGAGATTGTGATACAAACAGTAGA	883
Qy	956	GAGGAATGCCCTTC	969
Db	884	GACAGGGGCTTC	897

Search completed: September 9, 2005, 11:35:54
Job time : 5333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 9, 2005, 08:53:59 ; Search time 681 Seconds
(without alignments)
10579.038 Million cell updates/sec

Title: US-10-829-432-3
Perfect score: 1217
Sequence: 1 ggcgtcgcttcatcattcatca.....aaaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1217	100.0	1217	3	Aaz50160 Corn Aden
2	359.6	29.5	928	3	Aaz50163 Wheat Ade
3	350.2	28.8	378	12	Adj44283 Plant cDN
4	338.8	27.8	936	3	Aaz50162 Soybean A
5	320.4	26.3	627	3	Aac44691 Arabidops
6	316.2	26.0	890	3	Aaz50159 Corn Aden
7	296.4	24.4	1068	3	Aac35266 Arabidops
8	296.4	24.4	1175	3	Aac36588 Arabidops
9	296.4	24.4	1183	3	Aac45659 Arabidops
10	294.4	24.2	831	6	Abz12905 Arabidops
11	294.4	24.2	831	6	Adg87955 A. thalia
12	294.4	24.2	831	12	Adn74374 Thale cre
13	290	23.8	1130	3	Aac49070 Arabidops
14	288.4	23.7	1133	3	Aac36627 Arabidops
15	264.8	21.8	687	12	Adj39869 Plant cDN
16	263.6	21.7	917	3	Aaz98337 A. thalia
17	263.6	21.7	948	3	Adc42294 Arabidops
18	245	20.1	262	7	Adg66664 Corn seed
19	233.4	19.2	595	12	Adj43843 Plant cDN
20	225	18.5	714	3	Aac51190 Arabidops

21	198.2	16.3	483	3	AAC51315	Aac51315 Arabidops
22	197.2	16.2	431	3	Aaz50161	Aaz50161 Rice Aden
23	188.6	15.5	328	12	ADP92312	Adp92312 Cotton ex
24	179	14.7	597	8	ACA47863	ACA47863 Prokaryot
25	170.2	14.0	603	8	ACA27590	ACA27590 Prokaryot
26	168.6	13.9	642	8	ACA53672	ACA53672 Prokaryot
27	162.8	13.4	521	3	Aaz50164	Aaz50164 Wheat Ade
28	161.4	13.3	600	13	ADT42933	Adt42933 Bacterial
29	159.6	13.1	615	6	ABN92217	Abn92217 Staphyloc
30	159.6	13.1	615	13	ADS02000	Adg02000 Staphyloc
31	158.2	13.0	597	8	ACA46934	ACA46934 Prokaryot
32	156.4	12.9	609	13	ADS3268	Adg33268 Non-SCSSU
33	156.4	12.9	1160	3	AAA08597	Aaa08597 Saccharom
34	153.2	12.6	2014	10	ADD43524	Add43524 Human cDN
35	153.2	12.6	2014	12	ADK67851	Adk67851 Phosphoad
36	153.2	12.6	2014	12	ADK67854	Adk67854 Phosphoad
37	152.6	12.5	594	13	ADT48369	Adt48369 Bacterial
38	152.2	12.5	2424	12	ADK67852	Adk67852 Phosphoad
39	152	12.5	1845	10	ADD43532	Add43532 Humanopen
40	152	12.5	3774	8	ABX10363	Abx10363 DNA encod
41	152	12.5	3774	12	ADK67853	Adk67853 Phosphoad
42	152	12.5	3774	13	ADP23471	Adp23471 PRO poly
43	151.4	12.4	588	13	ADT41933	Adt41933 Bacterial
44	150.6	12.4	2617	3	Aaz94206	Aaz94206 Human tra
45	150.6	12.4	2617	8	ABX77649	Abx77649 Different

ALIGNMENTS

RESULT 1
ID AAZ50160 standard; cDNA; 1217 BP.
XX
AC AAZ50160;
XX
DT 04-MAY-2000 (first entry)
XX
DE Corn Adenylsulphate kinase-2 cDNA clone.
XX
KW Adenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW corn; clone p0016.ctscj40rb; transgenic plant; screen; antibody; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 2..1033
FT /*tag= a
FT /product= "Corn APS kinase-2"
FT /note= "Derived from clone p0016.ctscj40rb"
XX
PN WO200004165-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015809.
XX
PR 14-JUL-1998; 98US-0092833P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Falco SC, Allen SM, Anderson SL;
XX
DR WPI; 2000-182430/16.
XX
PT P-PSDB; AAY44789.
XX
PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'-
phosphosulphate kinase, useful for altering expression of sulfate
XX
XX assimilation protein in plants.
XX
XX Claim 3; Page 30; 42pp; English.

CC The present sequence is a cDNA encoding corn adenylsulphate kinase (APS
CC kinase), also known as adenosine-5' phosphosulphate kinase. This is
CC obtained from clone p0016.ctscj40rb, derived from corn pooled tassels
CC shoots, p0016 cDNA library. APS kinase is a sulphate assimilation
CC protein, that catalyses the conversion of adenosine-5' phosphosulphate
CC (APS) to 3'-phospho-adenosine-5' phosphosulphate (PAPS). This sequence is
CC used as probe and primers to identify, obtain and synthesise sulphate
CC assimilation proteins from other plants. It is also used to produce
CC transgenic plants, that are useful for altering the expression levels of
CC a sulphate assimilation protein. The APS kinase peptides are useful for
CC producing antibodies, that are used to screen and isolate cDNA clones
XX

SQ Sequence 1217 BP; 344 A; 292 C; 303 G; 278 T; 0 U; 0 Other;

Query Match 100.0%; Score 1217; DB 3; Length 1217;
Best Local Similarity 100.0%; Pred. No. 1.2e-297;
Matches 1217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTCCGTTTCATTTTCATCAATCAACAGAACCTCTGGTTCACACACGCGCAGCAACACC 60
DB 1 GCGTCGGTTTCATTTTCATCAATCAACAGAACCTCTGGTTCACACACGCGCAGCAACACC 60

QY 61 GAGCCAGCGCCGCGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 120
DB 61 GAGCCAGCGCCGCGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 120

QY 121 GAGCCAGCGCTCGCGCTCATCTCGTAAATCCACAGCGCGCTCGCGCTCGCGCTCGCG 180
DB 121 GAGCCAGCGCTCGCGCTCATCTCGTAAATCCACAGCGCGCTCGCGCTCGCGCTCGCG 180

QY 181 CCTCACCCCTAGCGATGCGCCACTCCCGCGCTCGTGATCCATGCGCTCACTCCCGCTTC 240
DB 181 CCTCACCCCTAGCGATGCGCCACTCCCGCGCTCGTGATCCATGCGCTCACTCCCGCTTC 240

QY 241 CTCACACTCTTCGCGGGTCTCGCCAGTGATAGTGGCGCCCGAGGGGAGGGCGCGGG 300
DB 241 CTCACACTCTTCGCGGGTCTCGCCAGTGATAGTGGCGCCCGAGGGGAGGGCGCGGG 300

QY 301 TCGCGGTAGCGACTCCACCGCGCATTTGGGCGGTGGGTGGCGCGCGCGCGGAATGG 360
DB 301 TCGCGGTAGCGACTCCACCGCGCATTTGGGCGGTGGGTGGCGCGCGCGCGGAATGG 360

QY 361 AGCAGCGCCGCGGGAGGCGCCGACAGCCAGTGAAGGAGAGCCCTGTAATGTCGAACAT 420
DB 361 AGCAGCGCCGCGGGAGGCGCCGACAGCCAGTGAAGGAGAGCCCTGTAATGTCGAACAT 420

QY 421 TGGGAAATCGACTAATATTTATGSCACAAATTCCTTGATGAGCAATCTGATAGACAGAA 480
DB 421 TGGGAAATCGACTAATATTTATGSCACAAATTCCTTGATGAGCAATCTGATAGACAGAA 480

QY 481 ATTGCTGGGACAAAAGGCTGTCTGTATGATTAACAGGACTCAGTGGTTTACGGGAAAAG 540
DB 481 ATTGCTGGGACAAAAGGCTGTCTGTATGATTAACAGGACTCAGTGGTTTACGGGAAAAG 540

QY 541 TACTCTTGATGTCAGTCTGAGTCTGAGTTCATTTGACAGGCCACCTCAGTATGTACT 600
DB 541 TACTCTTGATGTCAGTCTGAGTCTGAGTTCATTTGACAGGCCACCTCAGTATGTACT 600

QY 601 TGATGGTGCACCTCAGACATGGCTTAATAGATGATTTAAGCTTTAAGGCGAAGACCG 660
DB 601 TGATGGTGCACCTCAGACATGGCTTAATAGATGATTTAAGCTTTAAGGCGAAGACCG 660

QY 661 TGCAGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGTGATGCTGGTGTCTAT 720
DB 661 TGCAGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGTGATGCTGGTGTCTAT 720

QY 721 ATGCAATGTCAGCTGATATCTCCATACAGGAGAGATCGTGATGCATGCCGCTCTACT 780
DB 721 ATGCAATGTCAGCTGATATCTCCATACAGGAGAGATCGTGATGCATGCCGCTCTACT 780

QY 781 TCCACATCTCACTTATTGAAGTATTTATTCATTTGCCCTTAAATAATTTGGAAGCTCG 840
DB 781 TCCACATCTCACTTATTGAAGTATTTATTCATTTGCCCTTAAATAATTTGGAAGCTCG 840

QY 841 TGATCTTAAAGCCCTATACAGCTTGCACGTACAGGAAAGATTAAAGGTTTCACTGGAAT 900
DB 841 TGATCTTAAAGCCCTATACAGCTTGCACGTACAGGAAAGATTAAAGGTTTCACTGGAAT 900

QY 901 TGATGATCCATACGAAACCAACCAATTAATGGTGAAGATAGTAATTAAGATGAAGATGAGGA 960
DB 901 TGATGATCCATACGAAACCAACCAATTAATGGTGAAGATAGTAATTAAGATGAAGATGAGGA 960

QY 961 ATGCCCCCTCACCCAAAGCAATGGCCAGCAAGTTCATGCTACCTTGAAGAAAACGGATA 1020
DB 961 ATGCCCCCTCACCCAAAGCAATGGCCAGCAAGTTCATGCTACCTTGAAGAAAACGGATA 1020

QY 1021 TTTCGAGCTTAGTATATATGATTTTGAAGAGATGATCTGATTTCTGTGTCTCCATTACT 1080
DB 1021 TTTCGAGCTTAGTATATATGATTTTGAAGAGATGATCTGATTTCTGTGTCTCCATTACT 1080

QY 1081 TGTGGACACAATAAGATCTGTTGTTGGTTCACATGAATAAAAGGCATCAACATGTAGGAAG 1140
DB 1081 TGTGGACACAATAAGATCTGTTGTTGGTTCACATGAATAAAAGGCATCAACATGTAGGAAG 1140

QY 1141 TAAACAGAGGTACCGTTCATTCAGAAAAGGATATGGATTCGTTTAAAAAAGAAAAA 1200
DB 1141 TAAACAGAGGTACCGTTCATTCAGAAAAGGATATGGATTCGTTTAAAAAAGAAAAA 1200

QY 1201 AAAAAAAGAAAAA 1217
DB 1201 AAAAAAAGAAAAA 1217

RESULT 2
AAZ50163
ID AAZ50163 standard; cDNA; 928 BP.
XX AAZ50163;
AC AAZ50163;
XX
DT 04-MAY-2000 (first entry)
XX
DE Wheat Adenylsulphate kinase-1 cDNA clone.
XX
KW Adenylsulphate kinase; Adenosine-5'phosphosulphate kinase; APS kinase;
KW 3'-Phospho-adenosine-5'phosphosulphate; PAPS; sulphate assimilation;
KW wheat; clone wrl.pk0101.e2; transgenic plant; screen; antibody; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 3..743
FT /*tag= a
FT /product= "Wheat APS kinase-1"
FT /note= "Derived from clone wrl.pk0101.e2"
XX
FN WO200004165-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015809.
XX
PR 14-JUL-1998; 98US-0092833P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Falco SC, Allen SM, Anderson SL;
XX
DR WPI; 2000-182430/16.
XX P-PSDB; AAY44792.
XX
PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'
PT phosphosulfate kinase, useful for altering expression of sulfate
PT assimilation protein in plants.
XX
PS Claim 3; Page 34-35; 42pp; English.
XX

CC The present sequence is a cDNA encoding wheat adenylisulphate kinase
CC (APS kinase), also known as adenosine-5' phosphosulphate kinase. This is
CC obtained from clone wrl.pk101.e2, derived from 7 day old light grown
CC wheat root seedlings, wrl cDNA library. APS kinase is a sulphate
CC assimilation protein, that catalyses the conversion of adenosine-5'
CC phosphosulphate (APS) to 3'-phospho-adenosine-5' phosphosulphate (PAPS).
CC This sequence is used as a probe and primer to identify, obtain and
CC synthesise sulphate assimilation proteins from other plants. It is also
CC used to produce transgenic plants, that are useful for altering the
CC expression levels of a sulphate assimilation protein. The APS kinase
CC peptides are useful for producing antibodies, that are used to screen and
CC isolate cDNA clones

XX
SQ Sequence 928 BP; 285 A; 168 C; 250 G; 225 T; 0 U; 0 Other;

Query Match 29.5%; Score 359.6; DB 3; Length 928;

Best Local Similarity 65.5%; Pred. No. 1.4e-80;

Matches 558; Conservative 0; Mismatches 289; Indels 5; Gaps 2;

QY 367 GCCCGGGAGGCCCGCCAGCCAGCCAGTGAAGGAGAAAGCTGTAAATGTCGAACATTTGGGAA 426

DB 77 GCCCGTCAATGATCAGCCATGCGCAGGTATCGACAAGCTTGTGACCTCACTGTTGGGAA 136

QY 427 ATCGACTAATATTTATGCGCAATGCTTTGATTTGACAATCTGATAGACAGAAATGCT 486

DB 137 ATCGACAAACGTTCTTTGGCATGACTGTCCATAGTTCAGTTTGAGGCGCAGGAATGCT 196

QY 487 GGGACAAAAGCTGTGCTGATGATACAGACTCAGTGGTTCAGGAAAAGTACTCT 546

DB 197 AAATCAGAAGGGTGTGTTGTGTGATTAACAGGGTTAAGTGGTTCAGGAAAAGCAGCAT 256

QY 547 TGCATGTGCACTGAGTCGTGAGTTGCTATTCAGAGGCCACCTCAGCTATGATCTTGATGG 606

DB 257 AGCATGGCGCTAAGTCGGAGCTGCATCTCGAGAGTCACTGACTACATCTTAGACGG 316

QY 607 TCACAACCTCAGACATGCGCTAAATAGAGATTTAAGCTTTAAGGCGAGAACCGTGCAGA 666

DB 317 TGACAATCTAAGGCATGGGTAAACCGAGACCTCTGTTTCGAAGCAAAAGGACCGTCTGA 376

QY 667 AAATATACAGAGTGTGAGTGGCGAAGCTTTTTCGTGATGCTGTGTCATATGCAAT 726

DB 377 AAATATACGACAGTAGGAGAGTAGCAAAAGCTGTTTTCGAGATGCTGCTGATCTGCAT 436

QY 727 TCGTAGCTTGATATCTCCATACAGAGAGATCGTGATGATCGCCGTGCTACTTCCACA 786

DB 437 TCGTAGCTTGATATCACCTACAGAGTGAAGCGAGCGCTTGCCCAAAATTAATCCACA 496

QY 787 TTCTAATCTTATGAAGTATTTATTTGATTTGCCCTTAAATTTGTGAAGCTTCGTATCC 846

DB 497 TTCTAATCTTATGAAGTATTTTGAATGTGCCACTTGAAGTTTGTGAAGCTAGGATCC 556

QY 847 TAAAGCCTATACAGCTTGCAGTACAGGAAAGATTAAGTTTCACTGGAATTTGATGA 906

DB 557 AAAAGGCTTGTAAGCTTGGCCGTCGAGGAAAAATCAAGGGTTTACTGGAATTTGATGA 616

QY 907 TCCATACGAACCAACCAATTAATGCTGAGATAGTAATTAAGATGAAGATGAGGAATGCC 966

DB 617 TCCATATGAGGACCTTCTGACTGCGAGATAGTATACAGTGCAGGCTGTGATCTGGC 676

QY 967 TTCACCCAAAGCAATGGCCAGCAAGTTCTATGCTACCTTGAAGAAAAACGGATTTGCA 1026

DB 677 CAGCGCTAATCGATGGCTGATCAAGTTGTGTGCATATCTTGAAGCAATGAGTTCTTACA 736

QY 1027 AGCTTAG---TATATGATTTTGAAGAGTTGATCTGATCTTGTGTGTCCTACTTGT 1083

DB 737 GGAATAGAGACGTGCTATGATGAGAAAAACATTTCTGAATTTGATCGCCCAAGGATGT 796

QY 1084 GCACACAATPAAGATCTGTTGTTGGTTCACATGAATAA--AAGCATCAACATAGGAAGT 1141

DB 797 GAAATATGAGGTAGTATTTATCTCTAGAAAGAGTGAATAGTAGTAGAGAACATATATTT 856

QY 1142 AACAGAAGGTACGGTTTCATTCAGAAACGGATATGGATTCATTCGTTTAAAAAATAAAAA 1201

DB 1142 AACAGAAGGTACGGTTTCATTCAGAAACGGATATGGATTCATTCGTTTAAAAAATAAAAA 1201

DB 857 GACATAAAGATCGAATCTGTATCATTTAATAAATTTGAAATGTTTTCAGCAAAAAA 916

QY 1202 AAAAAAATAA 1213

DB 917 AAAAAAATAA 928

RESULT 3

ADJ44283/C

ID ADJ44283 standard; cDNA; 378 BP.

XX AC ADJ44283;

XX DT 06-MAY-2004 (first entry)

XX DE Plant cDNA #5283.

XX KW Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KW antifungal.

XX OS Eukaryota.

XX PN US2004016025-A1.

XX PD 22-JAN-2004.

XX PF 26-SEP-2002; 2002US-00260238.

XX PR 26-SEP-2001; 2001US-0325277P.

XX PR 26-SEP-2001; 2001US-0325448P.

XX PR 04-APR-2002; 2002US-0370620P.

XX PA (BUDW/) BUDWORTH P.

XX PA (MOUG/) MOUGHAMER T.

XX PA (BRIG/) BRIGGS S P.

XX PA (COOP/) COOPER B.

XX PA (GLAZ/) GLAZEBROOK J.

XX PA (GOFF/) GOFF S A.

XX PA (KATA/) KATAGIRI F.

XX PA (KREP/) KREPS J.

XX PA (PROV/) PROVART N.

XX PA (RICK/) RICHE D.

XX PA (ZHUT/) ZHU T.

XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;

XX Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;

XX WPI; 2004-190374/18.

XX New rice promoter, useful for manipulating crop plants to alter or

XX improve phenotypic characteristics, e.g. produce large quantities of oil

XX or proteins, resistance to insecticides, virus or fungi, stress tolerance

XX or high nutritional value.

XX Example 13; SEQ ID NO 5283; 230pp; English.

XX The invention relates to plant nucleotide sequences that direct seed-,
XX leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
XX or constitutive transcription of an operatively linked nucleic acid
XX segment. The invention also relates to a method for augmenting a plant
XX genome and a method of identifying a gene, where its expression is
XX altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
XX in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
XX canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
XX sorghum, rice or wheat. The polynucleotides and the polypeptides they
XX encode are useful for manipulating crop plants to alter or improve
XX phenotypic characteristics, to produce large quantities of oil or
XX proteins, to incur resistance to insecticides, viruses or fungi, and to
XX incur stress tolerance (e.g. salt, cold or drought) to ensure the plants

CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX
SQ Sequence 378 BP; 75 A; 125 C; 89 G; 89 T; 0 U; 0 Other;

```
Query Match      28.8%; Score 350.2; DB 12; Length 378;
Best Local Similarity 98.7%; Pred. No. 2.3e-78;
Matches 374; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy 245 CACTCTTCGCGGGTCTCCAGTCATAGTGGCGCCCGCGAGGGGAGGCGCGGGTGGC 304
Db |||||
Qy 378 CACTCTTCGCGGGTCTCCAGTCATAGTGGCGCCCGCGAGGGGAGGCGCGGGTGGC 319
Db |||||
Qy 305 CGTAGCAGTGCACGGGGGATTTGGCGGTGGTGGCGCGCGCGCGCGGAGTGGAC 364
Db |||||
Qy 318 CGTAGCAGTGCACGGGGGATTTGGCGGTGGTGGCGCGCGCGCGCGGAGTGGAC 259
Db |||||
Qy 365 GCGCC-CGCGGGAGGCGCGCAGCCAGTGAAGGAGAGCCTGTAATGTCGAACATTGG 423
Db |||||
Qy 258 GCGCCAGCGGGGATGCCCCGACAGCCAGTGAAGGAGAGCCTGTAATGTCGAACATTGG 199
Db |||||
Qy 424 GAAATCGACTAATATTTTATGGCACAATTCCTGTGATTCGACAACTCTGATAGACAGAAAT 483
Db |||||
Qy 198 GAAATCGACTAATATTTTATGGCACAATTCCTGTGATTCGACAACTCTGATAGACAGAAAT 139
Db |||||
Qy 484 GCTGGGACAAAAGCTGTGCTATGATTAACAGAGCTCAGTGGTTTCAGGGAAAAGTAC 543
Db |||||
Qy 138 GCTGGGACAAAAGCTGTGCTATGATTAACAGAGCTCAGTGGTTTCAGGGAAAAGTAC 79
Db |||||
Qy 544 TCTTCATGTGCACGTGAGTCGTGAGTTGCAATTCAGAGGCCACCTCAGTATGACTTGA 603
Db |||||
Qy 78 TCTTCATGTGCACGTGAGTCGTGAGTTGCAATTCAGAGGCCACCTCAGTATGACTTGA 20
Db |||||
Qy 604 TGGTGACAACTCAGACAT 622
Db |||||
Qy 19 TGGTGACAACTCAGACAT 1
```

RESULT 4
AAZ50162
ID AAZ50162 standard; cDNA; 936 BP.

XX AAZ50162;

XX 04-MAY-2000 (first entry)

XX Soybean Adenylsulphate kinase cDNA clone.

XX Adenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
KW 3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW soybean; clone sdp2c.pk013.all; transgenic plant; screen; antibody; ss.

OS Glycine max.

XX Key Location/Qualifiers
FH 1..936
FT CDS

FT /*tag= a
FT /product= "Soybean APS kinase"
FT /note= "Derived from clone sdp2c.pk013.all"

XX WO200004165-A1.

XX 27-JAN-2000.

XX 13-JUL-1999; 99WO-US015809.

XX 14-JUL-1998; 98US-0092833P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX

PI Falco SC, Allen SM, Anderson SL;
XX WPI; 2000-182430/16.
DR P-PSDB; AAY44791.
XX New nucleic acid molecule and chimeric gene encoding an adenosine-5'-
PT phosphosulphate kinase, useful for altering expression of sulfate
PT assimilation protein in plants.
XX Claim 3; Page 33; 42pp; English.

XX The present sequence is a cDNA encoding soybean adenylsulphate kinase
CC (APS kinase), also known as adenosine-5' phosphosulphate kinase. This is
CC obtained from clone sdp2c.pk013.all, derived from soybean developing
CC pods, sdp2c cDNA library. APS kinase is a sulphate assimilation protein,
CC that catalyses the conversion of adenosine-5' phosphosulphate (APS) to 3'-
CC -Phospho-adenosine-5' phosphosulphate (PAPS). This sequence is used as a
CC probe and primer to identify, obtain and synthesise sulphate assimilation
CC proteins from other plants. It is also used to produce transgenic plants,
CC that are useful for altering the expression levels of a sulphate
CC assimilation protein. The APS kinase peptides are useful for producing
CC antibodies, that are used to screen and isolate cDNA clones

XX Sequence 936 BP; 266 A; 206 C; 240 G; 224 T; 0 U; 0 Other;

```
Query Match      27.8%; Score 338.8; DB 3; Length 936;
Best Local Similarity 71.5%; Pred. No. 2.6e-75;
Matches 445; Conservative 0; Mismatches 177; Indels 0; Gaps 0;
```

```
Qy 410 ATGTCGAACATTGGGAATCGACTAATATTTATGGCACATTTGTTGATGGCAATCT 469
Db |||||
Qy 313 ATGTCGAATGTTGGGAATCGACAAACATTAATGTCGATGATTCCTCAATTCAGAAACA 372
Db |||||
Qy 470 GATAGACAGAATTTGCTGGGACAAAAGGCTGTGTCGTATGATGATAACAGGACTCAGTGGT 529
Db |||||
Qy 373 GATAGACAGAGCTGCTTCACCAACAGGCTGTATATGCTTAACATGGCTCAGCGGA 432
Db |||||
Qy 530 TCAGGGAAAAGTACTCTTGGCATGTGCATGAGTCGTGATGTTGCATTGCAGAGGCCACCTC 589
Db |||||
Qy 433 TCAGGGAAAAGCACTATTGTCATGTGCTCTGAGTCAAAAGCTTGCATCCAAAGGAAAAC 492
Db |||||
Qy 590 ACGTATGATCTTGTGATGGTGAACCTCAGACATGCGCTTAATAGAGATTTAAGCTTTAAG 649
Db |||||
Qy 493 TCTTACATCTCTGTGATGGTGAACATATTTCGGCATGCTCTAAACACAGGATCTTAGT 552
Db |||||
Qy 650 GCAGAAGACCTGCAGAAAATATACGAAGAGTTGGTGAAGTGCGAAAAGCTTTTTCCTCAT 709
Db |||||
Qy 553 GCAGAAGATCGTTCTGAAAACATTAGAAGGATTTGGTGAAGTGCGAAAAGCTTTTTCAGAT 612
Db |||||
Qy 710 GCTGGTGTATATGATCTAGCTTGTATATCTCCATACAGAGAGATCGTGTATGATGCG 769
Db |||||
Qy 613 GCTGGTGTATTTTGCATCACTAGTTTAATATACCATACCAAAAGGATAGAGATGATGCG 672
Db |||||
Qy 770 CGTGCTCTACTTCCACATCTTAACCTTTATTCGAATTTATTTGATTTTCCCTTAAATTT 829
Db |||||
Qy 673 AGAGCAGTACTTTCAAAGAGGAGATTTTATGAGGTTTTCATAGATGTTTCCACTACATG 732
Db |||||
Qy 830 TGTGAAGCTCGTGTATCTTAAAGGGCTATACAAAGCTTGCACGTACAGGAAAAGATTAAGGT 889
Db |||||
Qy 733 TGTGAAGCTAGGAGCCCAAGAGGACTCTACAAGCTTGTCTGAGCTGGAAGATCAAGGT 792
Db |||||
Qy 890 TTCACTGGAATGATGATCCATACGAACCAACCAATTAATGTTGATAGATGATTAAGATG 949
Db |||||
Qy 793 TTCCTGTTATAGATGATCCATATGAACCCCGTGTAGTTGTGTGATAGTATTACACAG 852
Db |||||
Qy 950 AAAGATGAGGAATGCCCTTCCACCAAGCAATGGCCAAAGCAAGTTCTATGCTTACCTTGA 1009
Db |||||
Qy 853 AAAGGAAGTGAAGTCTCCCAAGTATGGCTGAAGAGTGAATATCTTACTTGGAG 912
Db |||||
Qy 1010 GAAACCGGATATTTGCAAGCTT 1031
Db |||||
Qy 913 GAGAACGGATACCTGCGGGCTT 934
Db |||||
```


[illegible]

PT assimilation protein in plants.

XX
PS
XX Claim 3; Page 29; 42pp; English.

XX The present sequence is a cDNA encoding corn adenylylsulphate kinase (APS kinase), also known as adenosine-5' phosphosulphate kinase. This is obtained from clone cen3n.pk0088.b10, derived from corn endosperm, cen3n cDNA library. APS kinase is a sulphate assimilation protein, that catalyses the conversion of adenosine-5' phosphosulphate (APS) to 3'-phospho-adenosine-5' phosphosulphate (PAPS). This sequence is used as probes and primers to identify, obtain and synthesise sulphate assimilation proteins from other plants. It is also used to produce transgenic plants, that are useful for altering the expression levels of a sulphate assimilation protein. The APS kinase peptides are useful for producing antibodies, that are used to screen and isolate cDNA clones

XX
SQ Sequence 890 BP; 232 A; 239 C; 262 G; 157 T; 0 U; 0 Other;

Query Match 26.0%; Score 316.2; DB 3; Length 890;

Best Local Similarity 66.9%; Pred. No. 1.3e-69;

Matches 450; Conservative 0; Mismatches 223; Indels 0; Gaps 0;

QY 361 AGCAGCGCGGGAGGCGCCGACAGCCAGTGAAGGAGAGCCTGTAATGTCGAACAT 420

Db 5 AGCGCGCGCGCGCTGCGAGGATCAGCAGCAGCAGCGCGCTGGTGACCTGCACCGT 64

QY 421 TGGGAATCGCATATATTTATGCAATTTGTCATGATGACAACTCTGATAGACAGAA 480

Db 65 CGGGAATCGCAATCTCTGTGCAATGATGCGCATCGGGCAGAGGAGCGACAGGG 124

QY 481 ATTGCTGGGAAAAAGGCTGTGCTATGATGATAACAGGACTCAGTGGTTTCAGGAAAAAG 540

Db 125 TCTGCTGAACCAAGAGGCTGCTGTGTGATCATCTGSCCTAAGCGGTTTCAGGGAAG 184

QY 541 TACTCTTGATGTGACTGAGTCGTGATGATGATGATGATGATGATGATGATGATGATGAT 600

Db 185 CACGCTCGCGCTGCGCTGAGCGCGAGCTGCACGCGAGGCGCCACCTCACGTACGTCCT 244

QY 601 TGATGTTGACACCTCAGACATGGCTTAATAGATTAAGCTTTAAGGCGAGAGACCG 660

Db 245 CGACGGCGACACCTTCAGCAGCGGCTGACAGGACCTCAGCTTCGGAGCGAGGACCG 304

QY 661 TCGAAGAAATATACGAAGATTTGGTGAAGTGGCAAGCTTTTGTGCTGATGCTGGTGCAT 720

Db 305 CGCCGAGAACATCCGACAGATAGGGAAGTAGCGAAGCTGTTTCGCGCAGCTGGCTCGT 364

QY 721 ATGCATTGCTAGCTTGATATCTCCATACAGGAGATGATGATGATGATGATGATGATGAT 780

Db 365 CTGCATCGCCAGCTCATATCCCTCTACAGAGCGACCGCGCTGTCGCGATCTGCT 424

QY 781 TCCACATTCTAATTTTATTTGAAGTATTTATTTGATTTGCCCTTAAATTTTGAAGCTCG 840

Db 425 GCCAAGCACTGTTTATCGAGTGTCTTGAAGTGTGCGCTTCAAGTGTGCGAAGCCAG 484

QY 841 TGATCCTTAAAGCCCTATCAAGCTTTCAGCTACAGGAAAGATTAAGGTTTCACTGGAAAT 900

Db 485 GGACCCCAAGGCTCTCAAGCTCGCAGCGCGCGCAAAATCAAGGTTTCAAGGCTAT 544

QY 901 TGATGATCATGACCAACCAATTAATGATGATGATGATGATGATGATGATGATGATGATG 960

Db 545 CGACGATCCTTACGAACCGCGCTCGGAGCTGTGATGATGATGATGATGATGATGATGATG 604

QY 961 ATGCGCTTCCACCAAGCAATGGCCAGCAAGTCTATGCTACCTTGAAGAAACGGATA 1020

Db 605 CTGCGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 664

QY 1021 TTTGCAAGCTTAG 1033

Db 665 CTTCCAGGACTAG 677

RESULT 7
AAC35266

ID AAC35266 standard; DNA; 1068 BP.

XX
AC AAC35266;XX
DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 9570.

XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX
PN EP1033405-A2.XX
PD 06-SEP-2000.XX
PF 25-FEB-2000; 2000EP-00301439.XX
PR 05-FEB-1999; 99US-0121825P.XX
PR 05-MAR-1999; 99US-01231180P.XX
PR 09-MAR-1999; 99US-0123548P.XX
PR 23-MAR-1999; 99US-0125788P.XX
PR 25-MAR-1999; 99US-0126264P.XX
PR 29-MAR-1999; 99US-0126785P.XX
PR 01-APR-1999; 99US-0127462P.XX
PR 06-APR-1999; 99US-0128234P.XX
PR 08-APR-1999; 99US-0128714P.XX
PR 16-APR-1999; 99US-0129845P.XX
PR 19-APR-1999; 99US-0130077P.XX
PR 21-APR-1999; 99US-0130449P.XX
PR 23-APR-1999; 99US-0130510P.XX
PR 23-APR-1999; 99US-0130891P.XX
PR 28-APR-1999; 99US-0131449P.XX
PR 30-APR-1999; 99US-0132048P.XX
PR 30-APR-1999; 99US-0132407P.XX
PR 04-MAY-1999; 99US-0132484P.XX
PR 05-MAY-1999; 99US-0132485P.XX
PR 06-MAY-1999; 99US-0132486P.XX
PR 06-MAY-1999; 99US-0132877P.XX
PR 07-MAY-1999; 99US-0132863P.XX
PR 11-MAY-1999; 99US-0134256P.XX
PR 14-MAY-1999; 99US-0134218P.XX
PR 14-MAY-1999; 99US-0134219P.XX
PR 14-MAY-1999; 99US-0134221P.XX
PR 14-MAY-1999; 99US-0134370P.XX
PR 18-MAY-1999; 99US-0134768P.XX
PR 19-MAY-1999; 99US-0134941P.XX
PR 20-MAY-1999; 99US-0135124P.XX
PR 21-MAY-1999; 99US-0135353P.XX
PR 24-MAY-1999; 99US-0135629P.XX
PR 25-MAY-1999; 99US-0136021P.XX
PR 27-MAY-1999; 99US-0136392P.XX
PR 28-MAY-1999; 99US-0136782P.XX
PR 01-JUN-1999; 99US-0137222P.XX
PR 03-JUN-1999; 99US-0137528P.XX
PR 04-JUN-1999; 99US-0137502P.XX
PR 07-JUN-1999; 99US-0137724P.XX
PR 08-JUN-1999; 99US-0138094P.XX
PR 10-JUN-1999; 99US-0138540P.XX
PR 10-JUN-1999; 99US-0138847P.XX
PR 14-JUN-1999; 99US-0139119P.XX
PR 16-JUN-1999; 99US-0139452P.XX
PR 16-JUN-1999; 99US-0139453P.XX
PR 17-JUN-1999; 99US-0139492P.XX
PR 18-JUN-1999; 99US-0139454P.XX
PR 18-JUN-1999; 99US-0139455P.XX
PR 18-JUN-1999; 99US-0139456P.XX
PR 18-JUN-1999; 99US-0139457P.XX
PR 18-JUN-1999; 99US-0139458P.XX
PR 18-JUN-1999; 99US-0139459P.XX
PR 18-JUN-1999; 99US-0139460P.

PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145115P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0145386P.
PR	02-AUG-1999;	99US-0145388P.
PR	02-AUG-1999;	99US-0145389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	18-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 24.4%; Score 296.4; DB 3; Length 1068;		
Best Local Similarity 68.0%; Pred. No. 1.5e-64;		
Matches 430; Conservative 0; Mismatches 196; Indels 6; Gaps 1;		
Qy	411	TGTCGAACATTGGGAATCGACTAAATATTTATTTATGGCACAATGCTTGATTGGACAATCTG 470
Db	308	TGTCGACGTCGGAACCTCGACAAATATAAAGTGGCATGAATGTTCTGTGGAAAAGTTG 367
Qy	471	ATAGACAGAAATTCGTGGCAAAAAGCTGTGCTATGATTAACAGGACTCAGTGGTT 530
Db	368	ATAGACAGAGATTGCTTGATCAGAAAGATGTGATTGGGTACCGGTCTTAGTGGTT 427
Qy	531	CAGGAAAAGTACTCTTTGCATGTGCACGTGCTGAGTTCGTTCATTCGAGAGCCACCTCA 590
Db	428	CAGGGAAGAGTACTTTGCTTGTGTTGAATCAGATGTTCTATCAAAAGGGGAAGCTTT 487
Qy	591	CGTATGTACTTGTGTTGAGCAACCTCAGACATGCGCTAAATAGAGATTTAAGCTTTAAGG 650

XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-01218025P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 01-APR-1999; 99US-0126785P.
PR 06-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0128845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 28-APR-1999; 99US-0130510P.
PR 30-APR-1999; 99US-0130891P.
PR 04-MAY-1999; 99US-0131449P.
PR 05-MAY-1999; 99US-0132048P.
PR 06-MAY-1999; 99US-0132407P.
PR 07-MAY-1999; 99US-0132486P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140695P.
PR 29-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 09-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142920P.
PR 13-JUL-1999; 99US-0142977P.
PR 14-JUL-1999; 99US-0143542P.
PR 15-JUL-1999; 99US-0143624P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 27-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.

Db 515 CTGGAATAATCTGCATTGCGAGTTTGCATATCTCTTATAGAACAGATAGGACGCTTGTC 574
Qy 771 GTGCTCTACTCCACATCTTAATCTTATTGAAGTATTTATGATTTGCCCTTAAAAATTT 830
Db 575 GAAGTTTGTCTCCCGAGGAGATTTTGTGAGGTGTTCATGGATGACCGCTTAGTGTTT 634
Qy 831 GTGAAGCTCGTGATCTCTAAAGCCCTATACAAGCTTGCACGTACAGGAAGATTTAAAGGTT 890
Db 635 GCGAGCGAGGATCCAAAGGCTCTTACAAGCTTGTCTGTCGAGGAAGATCAAGGTT 694
Qy 891 TCATCGGAATGATGATCATACGACCAACCAATTAATTAAGTGTGAGATGAATTAAGATGA 950
Db 695 TTACCGGGATCGATGACCTTACGAGCCACCATTCGAATGCGAG-----ATTTCCTAG 748
Qy 951 AGATGAGGAAGTCCCTTACCAACCAAGCAATGGCCCAAGTCTTATGCTACCTTGAAG 1010
Db 749 GACGTGAAGGAGAACTTCTCTATCGAAATGGCGAAAGGTGTCGATACTTAGATA 808
Qy 1011 AAAACGGATATTTGCAAGCTTA 1032
Db 809 ACAAGGTTATCTTCAAGCATA 830

RESULT 12

ADN74374

ID ADN74374 standard; cDNA; 831 BP.

XX AC ADN74374;

XX DT 15-JUL-2004 (first entry)

XX DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2269.

XX KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;

KW growth regulator; animal feed product; thale cress;

KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

XX OS Arabidopsis thaliana.

XX XX WO2004035798-A2.

XX XX 29-APR-2004.

XX PF 20-OCT-2003; 2003WO-BP011658.

XX PR 18-OCT-2002; 2002EP-00079408.

XX PA (CROP-) CROPDESIGN NV.

XX PI Inze D, De Veylder L, Vlieghe K;

XX DR WPI: 2004-348466/32.

DR P-PSDB; ADN74375.

XX Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
PT more proteins.

XX PS Claim 1; SEQ ID NO 2269; 134pp; English.

XX CC This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreduplication, biochemistry, signal
CC transduction, storage lipid mobilisation and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these

CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polynucleotide sequence is thale cress cDNA
CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa
CC transcription factor, given in an exemplification of the invention.

XX SQ Sequence 831 BP; 224 A; 137 C; 223 G; 247 T; 0 U; 0 Other;

Qy Query Match 24.2%; Score 294.4; DB 12; Length 831;

Best Local Similarity 68.3%; Pred. No. 4.3e-64;

Matches 425; Conservative 0; Mismatches 191; Indels 6; Gaps 1;

Qy 411 TGTGCAATCTGGGAATCGACTAATATTTTATGACAAATTTGCTTGTATTGGACAATCTG 470

Db 215 TGTCTACCGTCGGAACCTCGACAAATATAAAGTGGCATGAATGTTCTGTTGAGAAAGTTG 274

Qy 471 ATAGACAGAAATTCGTGGGACAAAAAGGCTGTGTTATGGATAACAGGACTCAGTGTGT 530

Db 275 ATAGACAGAGATTGCTTGATCAGAAAGGATGCTGATTGGGTCAACGGTCTTAGTGTGT 334

Qy 531 CAGGAAAAGTACTCTTGCATGTGCATGCTGAGTTCGATTCGATTCGAGAGCCACCTCA 590

Db 335 CAGGGAAGAGTACTTGTGCTTTTGAATCAGATGTTGTATCAAAAGGGGAAGCTTT 394

Qy 591 CGTATGTACTTGTGATGGTGACAACTCGACATGGCTTAAATAGAGATTTAAAGTCTTAAGG 650

Db 395 GTTATATCTTGTGATGGTGATTAATGTAGGCATGCTTAAACCGTGATCTTAGCTTTAAG 454

Qy 651 CAGAAGCCGTGCAGAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGTCTGATG 710

Db 455 CTGAGGATCGTGCAGAGAATATTCGTAGATTGGAGAGTTGCTTAAGCTTTTGGCGATG 514

Qy 711 CTGTTGTCATATGATCTGCTAGCTTGATATCTCCATACAGGAGATCGTATCATGCC 770

Db 515 CTGGAATAATCTGCATTGCGAGTTTGATATCTCCTTATAGAACAGATAGGACGCTTGTG 574

Qy 771 GTGCTCTACTTCCACATTTAACTTTAATGAAGTATTTATTGATTTGCCCTTAAAAATTT 830

Db 575 GAAGTTTGTCTCCCGAGGGAGATTTTGTGAGTGTTCATGATGATCCGCTTAGTGTTT 634

Qy 831 GTGAAGCTCGTGATCTTAAAGCCCTATACAAGCTTGCACGTACAGGAAGATTTAAAGTTT 890

Db 635 GCGAGGCGAGGGATCCAAAGGGTCTTTACAAGCTTGTCTGTCGAGGAAGATCAAGGTTT 694

Qy 891 TCATCTGGAAATTTGATGATCCATACGAACCAACCAATTAATGTTGAGATAGTAAATGAAGA 950

Db 695 TTACCGGGATCGATGACCTTTACGAGCCACCATTTGAATCTGCGAG-----ATTTCCTAG 748

Qy 951 AAGATGAGGAATGCCCTTTACCCAAAGCAATGGCCCAAGATTTCTATGCTACCTTTGAAG 1010

Db 749 GACGTGAAGGAGGAACCTTCTCTATCGAAATGGCGGAAGAGTCTGTCGATACTTAGATA 808

Qy 1011 AAAACGGATATTTGCAAGCTTA 1032

Db 809 ACAAGGTTATCTTCAAGCATA 830

RESULT 13

AAC49070

ID AAC49070 standard; DNA; 1130 BP.

XX AC AAC49070;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 59815.

XX KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.

XX

OS Arabidopsis thaliana.
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301433.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-01213180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.

PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query Match 23.8%; Score 290; DB 3; Length 1130;
Best Local Similarity 70.2%; Pred. No. 6, 4e-63;
Matches 389; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY	416	AACATTGGGAATCGACTAATATTTATGGCACAATTGCTTGATGGACAATCTGATAGA	475
Db	339	ATCAATGGAAAGCAAAAGAACATTGCTGGCATGATTGTCGGTTACTAAATCCGACAGG	398
QY	476	CAGAATTCCTGGGCAAAAGGCTGTCGTATCGATAACAGACTCAGTGGTTCAGGG	535
Db	399	CAAGAATTAATTAAGCAAAAGGATGTGTGATTTGGATTACTGGCTTAAGTGTTTCAGGT	458
QY	536	AAAAGTACTCTTGCATGTGCATCGTGTGATTTGCATTTGCAGAGGCCACCTCACGTAT	595
Db	459	AAAAGTAGTCTGGCATGTGCTCTTAGTCGAGCTTTGCACATCGTGGAAAGCTTTCGTAT	518
QY	596	GTACTTGATGGTGACAACTCAGACATGGCCCTAAATAGAGATTTAAGCTTTAAGGCAGAA	655
Db	519	ATACTTGATGGTGACAACTTCCACATGGTTTAAACAGCGATCTTTAGTTTTCGAAGCAGAT	578
QY	656	GACCGTGCAAAATATACGAGAGTTGGTGAAGTGCGAAAGCTTTTTCGCTGATGCTGCT	715
Db	579	GATCAGCTGAAACATTCGAAGAGTTGGTGAAGTGGCTTAAACTGTTTGCAGATTCCTGCT	638
QY	716	GTCAATGCAATTCGTAGCTTGATCTATCTCCATACAGGAGATCGTGATGCGATCGCGTCT	775
Db	639	ATTATCTGATTGCAAGTTTATATCTCTTACCGGATAGACGAGCTGCGCTGCCGTGCA	698
QY	776	CTACTTCACATCTTAACCTTTTATGGAATTTATTTGATTTGCCCTTAAAAATTTGTGAA	835

Db	699	TTATTACCACAAGGAGATTTTCAITGAGGTATTTATGATGTGCCACTCCATGTTTGTGAA	758
QY	836	GCTCGTGTATCCTAAAGCCCTATACAAGTTTGCAGTACAGGAAGATTTAAAGGTTTCACT	895
Db	759	GCTAGAGATCCAAAAGGGCTTATACAAAACGTGTCACGCGCTGGTAAATCAAAGGTTTTCAC	818
QY	896	GGAATTGATGATCCATAGCAACCAACCAATTAATGTTGAGATAGTAGTAATTAAGATGAAAGAT	955
Db	819	GGAGTAGATGATCCATATGAAGCGCCTTTGGATTGCGAGATTGTAATACAAAAACAGTAGA	878
QY	956	GAGGAATGCCCTTC	969
Db	879	GACAAGGGGCTTTC	892
RESULT 14			
AAC36627			
ID	AAC36627 standard; DNA; 1133 BP.		
XX	AAC36627;		
XX	AAC36627;		
DT	17-OCT-2000 (first entry)		
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 14494.		
DE	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX	Arabidopsis thaliana.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
XX	06-SEP-2000.		
PD	25-FEB-2000; 2000EP-00301439.		
PF	25-FEB-1999;		
XX	99US-0121825P.		
XX	05-MAR-1999;		
PR	99US-0123180P.		
PR	09-MAR-1999;		
PR	99US-0123548P.		
PR	23-MAR-1999;		
PR	99US-0125788P.		
PR	25-MAR-1999;		
PR	99US-0126264P.		
PR	29-MAR-1999;		
PR	99US-0126785P.		
PR	01-APR-1999;		
PR	99US-0127462P.		
PR	06-APR-1999;		
PR	99US-0128234P.		
PR	08-APR-1999;		
PR	99US-0128714P.		
PR	16-APR-1999;		
PR	99US-0129845P.		
PR	19-APR-1999;		
PR	99US-0130077P.		
PR	21-APR-1999;		
PR	99US-0130449P.		
PR	23-APR-1999;		
PR	99US-0130510P.		
PR	28-APR-1999;		
PR	99US-0130891P.		
PR	30-APR-1999;		
PR	99US-0131449P.		
PR	30-APR-1999;		
PR	99US-0132048P.		
PR	04-MAY-1999;		
PR	99US-0132407P.		
PR	05-MAY-1999;		
PR	99US-0132485P.		
PR	06-MAY-1999;		
PR	99US-0132486P.		
PR	07-MAY-1999;		
PR	99US-0132487P.		
PR	11-MAY-1999;		
PR	99US-0132863P.		
PR	11-MAY-1999;		
PR	99US-0134256P.		
PR	14-MAY-1999;		
PR	99US-0134218P.		
PR	14-MAY-1999;		
PR	99US-0134219P.		
PR	14-MAY-1999;		
PR	99US-0134221P.		
PR	14-MAY-1999;		
PR	99US-0134370P.		
PR	18-MAY-1999;		
PR	99US-0134768P.		
PR	19-MAY-1999;		
PR	99US-0134941P.		
PR	20-MAY-1999;		
PR	99US-0135124P.		
PR	21-MAY-1999;		
PR	99US-0135353P.		
PR	24-MAY-1999;		
PR	99US-0135629P.		
PR	25-MAY-1999;		
PR	99US-0136021P.		
PR	27-MAY-1999;		
PR	99US-0136392P.		
PR	28-MAY-1999;		
PR	99US-0136782P.		
PR	01-JUN-1999;		
PR	99US-0137222P.		

PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0143977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144088P.
PR 16-JUL-1999; 99US-0144089P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match

23.7%; Score 288.4; DB 3; Length 1133;

Qy	757	TCGTGATGCATGCCGTGCTCTACTCCACATTCCTAACTTTATTGAAGTATTTATTGATTT	816
Db	365	TCGAAGCGCTGCCGCAAAATTACTACCCAATTCCTCGTTTCATTGAGGTGTTCTGGAATGT	424
Qy	817	GCCCCTAAAAATTTGTGAAGCTCGTGATCCTAAAGCCCTATACAAGCTTGCAGGTACAGG	876
Db	425	CCCACCTGAAGTATGTGAAGAAAGGGATCCAAAGGCCCTGTACAAGCTTGTCTCGTGCCGG	484
Qy	877	AAAGATTAAA-GGTTTCACTGGAATTGATGCATACGAACCAACCAATTAATGGTGAGA	935
Db	485	CAAAATCAAGGGCTTTACGGGAATAGATGATCCTTATGAAACACCTTCAGATTGTAG-	543
Qy	936	TAGTAATTAAGATGAAGAATGAGGAATGCCCTTACCCAAAGCAATGGCCAAAGCAAGTTC	995
Db	544	-----GGGGACTGGCCCTTCACCTAAATCAATGGCTGATCAAGTAG	583
Qy	996	TATGCTACCTTGAAGAAAGGATATTTGCAAGCTTAG	1033
Db	584	TGTCATATCTTGAAGCCAATGGATTTTCAGAACTAG	621

Search completed: September 9, 2005, 10:06:54
Job time : 686 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
1	159.6	13.1	615	3	US-09-134-001C-1680	Sequence 1680, Appl	
2	156.4	12.9	1160	3	US-09-153-310-41	Sequence 41, Appl	
3	153.2	12.6	2014	4	US-09-898-165B-1	Sequence 1, Appl	
4	152.2	12.5	2015	4	US-09-949-016-1925	Sequence 1925, Appl	
5	152	12.5	1845	4	US-09-898-165B-9	Sequence 9, Appl	
6	150.6	12.4	2617	4	US-09-786-240-21	Sequence 21, Appl	
C	7	146	12.0	640681	4	US-09-790-988-1	Sequence 1, Appl
8	139	11.4	1875	4	US-09-949-016-1468	Sequence 1468, Appl	
9	139	11.4	2506	1	US-08-879-561-4	Sequence 4, Appl	
10	132.8	10.9	1851	4	US-09-898-165B-10	Sequence 10, Appl	
11	132.8	10.9	2000	4	US-09-898-165B-2	Sequence 2, Appl	
12	128.4	10.6	681	4	US-09-543-681A-2395	Sequence 2395, Appl	
13	119.4	9.8	651	4	US-09-489-033A-2803	Sequence 2803, Appl	
14	94.2	7.7	89892	4	US-09-949-016-13667	Sequence 13667, A	
15	80.6	6.6	109925	4	US-09-949-016-13210	Sequence 13210, A	
C	16	79.2	6.5	601	3	US-09-949-016-50640	Sequence 50640, A
17	73.2	6.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl	
18	71.6	5.9	4403765	3	US-09-103-840A-2	Sequence 2, Appl	
19	48.4	4.0	1141	4	US-09-806-708B-22	Sequence 22, Appl	
20	48.4	4.0	1959	4	US-09-252-991A-7983	Sequence 7983, Appl	
C	21	48.4	4.0	2010	4	US-09-252-991A-8169	Sequence 8169, Appl
22	47.8	3.9	7218	1	US-08-232-463-14	Sequence 14, Appl	
C	23	45.6	3.7	1323	3	US-09-413-452-3	Sequence 3, Appl
24	45.6	3.7	1323	3	US-09-413-068-3	Sequence 3, Appl	
25	45.6	3.7	1323	3	US-09-403-345A-3	Sequence 3, Appl	
26	45.4	3.7	98844	3	US-09-791-211-10	Sequence 10, Appl	
27	45.4	3.7	143776	4	US-09-949-001-29	Sequence 29, Appl	

Qy	492	AAA-----AAGCTGTGTCGTATGATAACAGAGACTCAGTGGTTCAGGGAAA	539
Db	187	CAAGGGGTGGTTCGAGGATGACCGGTGTGGCTAAACAGGTCTCTCTGTGGTCTCGAAAA	246
Qy	540	GTACTCTTCGCATGTGCACCTGAGTCGTGAGTTGCAATTCAGAGGCCACCTCACGTATGTAC	599
Db	247	CAACGATAAGTTTGGCCCTGGAGGAGTACCTTGCTGCCATGCCATCCCTTGTTTACTCCC	306
Qy	600	TTGATGGTGCACACCTTCAGACATGCGCTTAAATAGAGATTTTAAGCTTTTAAGCGCAAGAC	659
Db	307	TGATGGGCAACAATGCTCCGTATGCGCTTAACAGAAATCTCGGATCTCTCTCTGGGACA	366
Qy	660	GTGCAGAAATATACAGAGATTGTTGAAGTGGCAAGACTTTTTCGTGATGCTGGTGCTCA	719
Db	367	GAGAGGAAATATCCGCCGAGATTGCTGAGGTGGCTAAAGCTGTTTGTGATGATGCTGGCTG	426
Qy	720	TATGCATTGCTAGCTTGATATCTCCATACAGAGAGATCGTGATGCATGCCGTGCTCTAC	779
Db	427	TCGCATTAACAGCTTTATTTCTCATTCGCAAGAGATCGTGAGATGCCCGCAAAATAC	486
Qy	780	TT-----CCACATTTAACTTTATTGGAAGTATTTATTGATTTGCCCCCTAAAAAATTTGTG	833
Db	487	ATGAATCAGCAGGGTGCCATCTTTTGAATAATTTGTAGATGCACCTCTAAATAATTTGTG	546
Qy	834	AAGCTCGTGATCCTTAAAGCCCTATACAAGCTTGCACGTACAGAAAGATTAAAGTTTCA	893
Db	547	AAAGCAGAGACGTAAGAGCCCTCTATAAAGGGCCAGAGCTGGGGAGATTAAAGGATTTA	606
Qy	894	CTGGAATTGATGATCCATACGAACCACTAATTAATGTTGAGATAGTAATTAAGATGAA	951
Db	607	CAGGTATTGATTCTGATTATGAGAAACCTGAAACTCCTGAGCGTGTCTTAAAAACCA	664

RESULT 4

US-09-949-016-1925

; Sequence 1925, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1925

; LENGTH: 2015

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-1925

Query Match 12.5%; Score 152.2; DB 4; Length 2015;

Best Local Similarity 59.7%; Pred. No. 4e-31;

Matches 276; Conservative 0; Mismatches 180; Indels 6; Gaps 1;

Qy	496	AGCGTGTGTCGTATGATAACAGAGACTCAGTGGTTCAGGGAAAGTACTCTTCATGTGC	555
Db	204	AGATGTACCGTGGCTTAAACAGGTCTCTCTGGTGTGGAAAAACACGATAAGTTTTCG	263
Qy	556	ACTGAGTCGTGAGTTGCAATTCGAGAGGCCACCTCACGTATGTACTTGAATGGTGACACCT	615
Db	264	CCTGGAGGAGTACCTTGTCTCCATGCCATCCCTTGTTACTCCCTGATGGGACATGT	323
Qy	616	CAGACATGCGCTTAAATAGAGATTTTAAGCTTTAAGCGCAGAGACCGTGCAGAAAAATATCG	675
Db	324	CCGTATGCGCTTAAACAGAAATCTCGAATCTCTCTCTGGGACAGAGAGGAAAAATATCG	383

Db 423 GCCATTCTTTGAAATATTTGTAGATGCACCTCTAAATATTTGTGAAAGCAGAGACGTAAA 482
QY 850 AGGCTTATCAAGCTTGCACGTACAGGAAAGATTAAAGTTTTCACCTGGAATTTGATGCC 909
Db 483 AGGCTCTATATAAAGGCCAGAGCTGGGGAGATTAAAGATTACAGGATTTGATCTGA 542
QY 910 ATACGAACCAACCAATTAATGTTGAGATAGTAATTAAGATGAA 951
Db 543 TTATGAGAAACCTGAAACCTCCTGAGCGTGTCTTAAACCAA 584

RESULT 6

US-09-786-240-21
; Sequence 21, Application US/09786240
; Patent No. 6558935
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PF-0592 PCT
; CURRENT APPLICATION NUMBER: US/09/786,240
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 60/133,642
; PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PERL Program
; SEQ ID NO 21
; LENGTH: 2617
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6558935 1420940CB1
US-09-786-240-21

Query Match 12.4%; Score 150.6; DB 4; Length 2617;
Best Local Similarity 59.5%; Pred. No. 1.3e-30;
Matches 275; Conservative 0; Mismatches 181; Indels 6; Gaps 1;
QY 496 AGGCTGTGTCGATGATGAATCAAGGACTCAGTGGTTTCAGGAAAAAGTACTCTTGCATGTGC 555
Db 215 AGGATGTACCGTGGCTAACAGGCTCTCTCGTGTGCTGGAAAAACAACGATAAGTTTTCG 274
QY 556 ACTGAGTCGTGAGTTGCATTGCAGAGGCCACCTCAGTGATGACTTGTGTTGTTGACAACT 615
Db 275 CCTGAGGAGTACCTTGTCTCCATGCCATCCCTTGTACTCCCTGGATGGGGAACAATGT 334
QY 616 CAGACATGCCCTAAATAGAGATTTAAGCTTTAAGGCAGAGACCCTGCAGAAAAATATACG 675
Db 335 CCGTCATGSCCTTAACAGAAATCTCGATCCTCTCTCGGGACAGAGAGAAAAATATCG 394
QY 676 AAGAGTTGTGAGTGGCAAGCTTTTTCGATGCTGTGTGTCATATGCAATTTGCTAGCTT 735
Db 395 CCGGATTTGTGAGTGGCTTAAGCTGTTTTGTGATGCTGTGCTGTGATTTACCAGCTT 454
QY 736 GATATCTCCATACAGAGAGATCGTGATGCATGCCGTGCTCTACTT-----CCACATTC 789
Db 455 TATTTCTCCTATTCGCAAGGATCGTGAGATGCCGCCCAAAATACATGATCAGAGGGCT 514
QY 790 TAACTTTTATGAAGTATTTTATTTGATTTGCCCTTAAATAATTTGTGAAGCTCGTGATCTTAA 849
Db 515 GCCATTTCTTTGAAATATTTGTAGATGCACCTCTAAATATTTGTGAAAGCAGAGACGTAAA 574
QY 850 AGGCTTATCAAGCTTGCACGTACAGGAAAGATTAAAGTTTTCACCTGGAATTTGATGCC 909
Db 575 AGGCTCTATAAAGGGCCAGAGCTGGGGAGATTAAAGGATTTTACAGGTTATTTGATCTGA 634

QY 910 ATACGAACCAACCAATTAATGTTGAGATAGTAATTAAGATGAA 951
Db 635 TTATGAGAAACCTGAAACCTCTGAGCGTGTCTTAAACCAA 676

RESULT 7

US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 12.0%; Score 146; DB 4; Length 640681;
Best Local Similarity 55.0%; Pred. No. 4.9e-28;
Matches 287; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

QY 434 AATATTTTATCGCACAAATTTGCTTGGACAATCTGTAGACAGAAATTTCTCGGACAA 493
Db 457777 AATATTTTGGCAAAAACATTTCAATTACACGCCTTAAACCTGAAAAAATTTGGTCTAT 457718
QY 494 AAAGGCTGTGTCGTATGATAACAGGACTCAGTGGTTTCAGGAAAAAGTACTCTTTGCATGT 553
Db 457717 AAATCAATTTGACTATGTTCTCCTGGGCTCTCAGGTTTCAGGAAAAATCAACCATAGCAAT 457658
QY 554 GCATGAGTCGTGAGTTGCATTGCAGAGGCCACCTCAGCTATGTACTTGTATGTTGAGAAC 613
Db 457657 TTTTATAGAGAAATATTTGTTTAAAAATGSAATTAATAGCTATTTATTAGACGGAGACAAT 457598
QY 614 CTCAGACATGCGCTAAATAGAGATTTAAGCTTTAAGGCAGAGACCGTGCAGAAAAATATA 673
Db 457597 ATTAGTCTGGTTTATGTTCTGATTTAAGTTTTAGTTTTTGGCAGATAGGAATGAAAAATTT 457538
QY 674 CGAAGAGTTGCTGAAGTGGCAAAAGCTTTTTCGCTGATGCTGCTGCATATGCAATTTGCTAGC 733
Db 457537 AGACGCAATGGAGAAGTAGTAAATGATGTTACATGCTGCTCTAATAATATTGTTATCA 457478
QY 734 TTGATATCTCCATACAGAGAGATCGTGATGCGATGCGCTGCTCTACTTCCCAATTTCTTAAC 793
Db 457477 GTTATTTCCCTATAGAAATCAAAGGGAAATGGTACGCTCAATTTAGTATAGAAAAAAT 457418
QY 794 TTTTATGAAGTATTTATTTGATTTGCCCTTAAATTTTGTGAAGCTCGTGATCTCTTAAAGGC 853
Db 457417 TTTTATGAAGTTTTTCATTTGATACCACTATTGAAATTTGTGAATTTCCGAGATCTCTNAAAA 457358
QY 854 CTATACAGCTTGCACGTACAGGAAAGATTAAAGTTTTCTCTGGAATTTGATGATCCATAC 913
Db 457357 TTATATAACAAGCCCGCACAGGCCAAATATCTGATTTTTTACTGGTATACAAATGTACATAC 457298
QY 914 GAACCAACCAATTAATGTTGAGATAGTAATTAAGATGAAAGAT 955
Db 457297 GAAACTCTATAACACCTGATGTTCTTTTAAAGGGGTACAGAT 457256

RESULT 8

US-09-949-016-1468
; Sequence 1468, Application US/09949016

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1468
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1468

Query Match 11.4%; Score 139; DB 4; Length 1875;
Best Local Similarity 58.2%; Pred. No. 1.6e-27;
Matches 266; Conservative 0; Mismatches 185; Indels 6; Gaps 1;

QY 497 GCCTGTGTCGTATGATACAGGACTCAGTGGTTCAGGGAAAAGTACTCTTGCATGTGCA 556
DB 154 GTTTCACAGTTTGGCTAACAGCTTGTCTGAGCGGGAAGACTACTGTGAGCATGGCC 213
QY 557 CTGAGTCTGTCGTGTCATGACAGAGCCACCTCAGCTATGTACTTGTGATGGTGACAACTC 616
DB 214 TTGGAGGAGTACCTGCTTTTTCATGGTATTCATCTCAGCTCTGGATGGTGACAATATT 273
QY 617 AGACATGGCTAAATAGAGATTAACTTTAAGGACAGAGACCGTGCAGAAAATATACGA 676
DB 274 CGTCAAGTCTCAATAAAAACTTGGCTTTAGTCTCTGAAGACAGAGAGAAATGTTTCA 333
QY 677 AGAGTTGGTGAAGTGCAGAGCTTTTGTCTGATGCTGGTGTCTATGATGATCTAGCTTG 736
DB 334 CGCATCGCAGAGTTGCTTAACTGTTTGAGATGCTGGCTTAGTGTGATCATCAAGTTTC 393
QY 737 ATATCTCCATACAGGAGATCG-----TGATGCATGCCGTCTCTACTTCCACATTTCT 790
DB 394 ATATCACCTTACCTCAGGATCGCAACAAATGCAAGCAAAATTCATGAGGTCAAGTTTA 453
QY 791 AACTTTATTGAAGTATTATTGATTTGGCCCTTAAATAATTTGTGAAGCTCGTGATCCTAAA 850
DB 454 CCGTTTTTTTGAAGTATTTTGATGCTCCTCTGATGCTTTTGTGAACAGAGGGATGTCAA 513
QY 851 GGCCTATACAGCTTCGACGTACAGGAAGATTAAAGGTTTTCACCTGGAATTGATGATCCA 910
DB 514 GGACTCTACAAAAAGCCCGGACAGAGAAATTAAGGTTTTCACCTGGGATGATTTCTGAA 573
QY 911 TACGAACCAACCAATTAATGGTGAGATAGTAATTAAGA 947
DB 574 TATGAAAGCCAGAGGCCCTCGAGTTGGTGTCTGAAAA 610

RESULT 9

US-08-879-561-4
; Sequence 4, Application US/08879561
; Patent No. 5817482
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Hillman, Phillip R.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/08/879,561
; APPLICATION NUMBER: US/08/879,561
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0325 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2506 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LUNGNOT02
; CLONE: 373887
US-08-879-561-4

Query Match 11.4%; Score 139; DB 1; Length 2506;
Best Local Similarity 58.2%; Pred. No. 1.8e-27;
Matches 266; Conservative 0; Mismatches 185; Indels 6; Gaps 1;

QY 497 GCCTGTGTCGTATGATACAGGACTCAGTGGTTCAGGGAAAAGTACTCTTGCATGTGCA 556
DB 189 GGTTCACAGTTTGGCTAACAGCTTGTCTGAGCGGGAAGACTACTGTGAGCATGGCC 248
QY 557 CTGAGTCTGTCGTGTCATGACAGAGCCACCTCAGCTATGTACTTGTGATGGTGACAACTC 616
DB 249 TTGAGAGGATACCTGCTTTTGTGATGTTTTCATGTTTTCATGCTACACTCTGGATGGTGACAATATT 308
QY 617 AGACATGGCTTAAATAGAGATTAAAGCTTTAAGGCAGAGACCGTGCAGAAAATATACGA 676
DB 309 CGTCAAGTCTCAATAAAAACTTGGCTTTAGTCTCTGAAGACAGAGAGAAATGTTTCA 368
QY 677 AGAGTTGGTGAAGTGGCAAGCTTTTGTCTGATGCTGTGTCATATGATGATGCTAGCTTG 736
DB 369 CGCATCGCAGAGTTGTCTTAACTGTTTGCAGATGCTGGCTTAGTGTGATCACAAGTTTC 428
QY 737 ATATCTCCATACAGGAGATCG-----TGATGCATGCCGTCTCTACTTCCACATTTCT 790
DB 429 ATATCACCTTACCTCAGGATCGCAACAAATGCAAGGCAAAATTCATGAGGTCAAGTTTA 488
QY 791 AACTTTATTGAAGTATTATTGATTTGGCCCTTAAATAATTTGTGAAGCTCGTGATCCTAAA 850
DB 489 CCGTTTTTTTGAAGTATTTTGATGCTCTCTGATGTTTGTGAACAGAGGGATGTCAAA 548
QY 851 GGCCTATACAGCTTCGACGTACAGGAAGATTAAAGTTCACCTGGAAATGATGATCCA 910
DB 549 GGACTCTACAAAAAGCCCGGACAGAGAAATTAAGGTTTTCACCTGGGATGATTTCTGAA 608
QY 911 TACGAACCAACCAATTAATGGTGAGATAGTAATTAAGA 947
DB 609 TATGAAAGCCAGAGGCCCTCGAGTTGGTGTCTGAAAA 645

```
RESULT 10
US-09-898-165B-10
; FILE REFERENCE: 18810-81552
; CURRENT APPLICATION NUMBER: US/09898165B
; PATENT NO. 6818428
; GENERAL INFORMATION:
; APPLICANT: Daniel H. Cohn
; APPLICANT: Muhammad Faiyaz ul Haque
; APPLICANT: Lily M. King
; APPLICANT: Deborah Krakow
; TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate
; TITLE OF INVENTION: (PAPS) Synthetic Enzymes and Methods for Treating
; TITLE OF INVENTION: Osteoarthritis Disorders
; FILE REFERENCE: 18810-81552
; CURRENT APPLICATION NUMBER: US/09/898,165B
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/399,212
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1851
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-898-165B-10

Query Match      10.9%; Score 132.8; DB 4; Length 1851;
Best Local Similarity 57.1%; Pred. No. 7.6e-26;
Matches 264; Conservative 0; Mismatches 192; Indels 6; Gaps 1;

QY 496 AGGCTGTGTCGATGATGAATACAGGACTCAGTGGTTTCAGGGAAAAGTACTCTTGCATGTGC 555
Db 126 AGGATGTACCGTGTGGCTAACAGGTCTCTCTGGTCTGGGAAAACCAACATAAGCTTTGC 185
QY 556 ACTGAGTCGTGATTCGATTCAGAGGCCACTCAGTGTACTGTGATGTTGATGTTGACAACT 615
Db 186 TTGGAAGAGTACCTTGTATCTCAGCCATCCATGTTACTCCCTGGATGGGCAATGT 245
QY 616 CAGCATGGCCCTAAATAGAGATTTAAAGCTTTAAAGGCAGAGACCGTGCAGAAAAATATACG 675
Db 246 CCGTCATGGCCCTTAATAGAACTGGGATTTCTCGCGGGACCGAGAGAGATATCCG 305
QY 676 AAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATGCTGGTGTCTATGATGATGCTAGCTT 735
Db 306 CCGGATCGGAGGTGGCAAGCTTTTTCGCGACGCGCGCTGGTTTGCATCACCAGCTT 365
QY 736 GATATCTCCATACAGGAGATCGTGATCGATCGGTGCTCT-----ACTTCCACATTC 789
Db 366 TATCTCTCTTTTGCAGAGATCGTGATCGGATGCGCGGAAAAATCCACGAATCAGCAGGACT 425
QY 790 TAACTTTTATGAAGTATTTATTTGATTTGCCCTTAAATAATTTGTGAAGCTCGTGATCTTAA 849
Db 426 CCGCTTCTTTGAGATCTTTGATGATCGGCTTTTAAATATCTGTGAAGCCGAGAGCTAAA 485
QY 850 AGGCTATACAAAGCTTGACGTACAGGAAAGATTAAGGTTTCACTGGAATGTGATGCC 909
Db 486 AGGACTCTACAAACGAGCCGACGAGAGAGATTAAGGGTTTACAGGCATCGATCTTGA 545
QY 910 ATACGAACCAACCAATTAATGTTGATAGTAGTAATTAAGATGAA 951
Db 546 CTATGAGAAACCTGAAACTCCAGAGTGTGCTGAAGACCAA 587

RESULT 11
US-09-898-165B-2
; Sequence 2, Application US/09898165B
; Patent No. 6818428
; GENERAL INFORMATION:
; APPLICANT: Daniel H. Cohn
; APPLICANT: Muhammad Faiyaz ul Haque
; APPLICANT: Lily M. King
; APPLICANT: Deborah Krakow
; TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate
; TITLE OF INVENTION: (PAPS) Synthetase Proteins and Methods for Treating
```

```
; TITLE OF INVENTION: Osteoarthritis Disorders
; FILE REFERENCE: 18810-81552
; CURRENT APPLICATION NUMBER: US/09/898,165B
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/399,212
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-898-165B-2

Query Match      10.9%; Score 132.8; DB 4; Length 2000;
Best Local Similarity 57.1%; Pred. No. 8e-26;
Matches 264; Conservative 0; Mismatches 192; Indels 6; Gaps 1;

QY 496 AGGCTGTGTCGATGATGAATACAGGACTCAGTGGTTTCAGGGAAAAGTACTCTTGCATGTGC 555
Db 186 AGGATGTACCGTGTGGCTAACAGGTCTCTCTGGTCTGGGAAAACCAACATAAGCTTTGC 245
QY 556 ACTGAGTCGTGATTCGATTCAGAGGCCACTCAGTGTACTGTGATGTTGATGTTGACAACT 615
Db 246 TTGGAAGAGTACCTTGTATCTCAGCCATCCCATGTTACTCCCTGGATGGGCAATGT 305
QY 616 CAGCATGGCCCTAAATAGAGATTTAAAGCTTTAAAGGCAGAGACCGTGCAGAAAAATATACG 675
Db 306 CCGTCATGGCCCTTAATAGAACTGGGATTTCTCGCGGGACCGAGAGAGATATCCG 365
QY 676 AAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATGCTGGTGTCTATGATGATGCTAGCTT 735
Db 366 CCGGATCGGAGGTGGCAAGCTTTTTCGCGACGCGCGCTGGTTTGCATCACCAGCTT 425
QY 736 GATATCTCCATACAGGAGATCGTGATCGATCGGTGCTCT-----ACTTCCACATTC 789
Db 426 TATCTCTCTTTTGCAGAGATCGTGATGCGGATGCGCGGAAAAATCCACGAATCAGCAGGACT 485
QY 790 TAACTTTTATGAAGTATTTATTTGATTTGCCCTTAAATAATTTGTGAAGCTCGTGATCTTAA 849
Db 486 CCGCTTCTTTGAGATCTTTGATGATCGGCTTTTAAATATCTGTGAAGCCGAGAGCTAAA 545
QY 850 AGGCTATACAAAGCTTGACGTACAGGAAAGATTAAGGTTTCACTGGAATGTGATGCC 909
Db 546 AGGACTCTACAAACGAGCCGACGAGAGAGATTAAGGGTTTACAGGCATCGATCTTGA 605
QY 910 ATACGAACCAACCAATTAATGTTGATAGTAGTAATTAAGATGAA 951
Db 606 CTATGAGAAACCTGAAACTCCAGAGTGTGCTGAAGACCAA 647

RESULT 12
US-09-543-681A-2995
; Sequence 2995, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2995
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2995

Query Match      10.6%; Score 128.4; DB 4; Length 681;
Best Local Similarity 55.3%; Pred. No. 7e-25;
```

Matches	280;	Conservative	0;	Mismatches	211;	Indels	15;	Gaps	1;
Qy	435	ATATTTTATGGCAAAATGCTTGATGGACAATCTGATAGACAGAAAATGCTGGACAAA	494						
Db	83	ATATTGCTGGCATCTCATCAAAATAGGTTTAAAAAGAGCGTGAAGCACAACAGGTACACA	142						
Qy	495	AAGCTCTGTCGTATGATACAGACTCAGTGGTTCAGGGAAAAAGTACTCTTGCATGTG	554						
Db	143	AAGGATGTGTACTTTGGTTTACTCGGTTATCTGGGTCAAGTAAATCAACACTGGCTGATG	202						
Qy	555	CACGTAGTCGTGAGTGTGCATTGCAGA-----GGCCACCTCACGTATGTAC	599						
Db	203	CGCTAGAGCAAACTTATATCAGTACTCGACACTCCATCGCCCTATCCGCACCTATTATT	262						
Qy	600	TTGATGTTGACAACTCAGACATGCGCCCTAAATAGAGATTTTAAGCTTTAAGGCAGAAAGC	659						
Db	263	TAGATGTTGATAATCTACGCCATGGTTTATGCCATGATCTTTGGGTTAGTGAACAAGATA	322						
Qy	660	GTGCAGAAAATATACGAAGAGTGTGTAAGTGGCAAGCTTTTTCGTGATGCTGGTGTC	719						
Db	323	GGCATGAAAAATATTTCGCCGTGTAGGGGAAGTGGCTAAATTAATGGTTGATGCCGATTAA	382						
Qy	720	TATGCATTTGCTAGCTTGATATCTCCATACAGAGAGATCGTGATGCGCTGCTCTAC	779						
Db	383	TTGTCTTAAACAGCAATTTATTTCTCTTATCAGAAATAGACACAAGATAAGAAAGGT	442						
Qy	780	TTCCACATTTCTAACTTTATTTGAAGTATTTATTTGATTTTGCCTCTAAAAATTTGTGAAGCTC	839						
Db	443	TTGCTCAAGGGCGATTATTAGATCTTTTGTGATACACCTTTAGCCCTTTGTGAAGCAC	502						
Qy	840	GTGATCTTAAAGGCTTATACAAAGCTTGCACTGACAGGAAAGATTAAAGTTTCACTGAA	899						
Db	503	GTGATCTTAAAGGCTCTTATCAAAAAAGCGCAGGAGAGATCAAAACAGTTTTTCGGCA	562						
Qy	900	TTGATGATCCATACGAACCCCAATT	925						
Db	563	TTGATTCACCTTTATGAACCCCCCACT	588						
RESULT 13									
US-09-489-039A-2803									
; Sequence 2803, Application US/09489039A									
; Patent No. 6610836									
; GENERAL INFORMATION:									
; APPLICANT: Gary Breton et. al									
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA									
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS									
; FILE REFERENCE: 2709.2004001									
; CURRENT APPLICATION NUMBER: US/09/489,039A									
; PRIOR FILING DATE: 2000-01-27									
; PRIOR APPLICATION NUMBER: US 60/117,747									
; PRIOR FILING DATE: 1999-01-29									
; NUMBER OF SEQ ID NOS: 14342									
; SEQ ID NO 2803									
; LENGTH: 651									
; TYPE: DNA									
; ORGANISM: Klebsiella pneumoniae									
US-09-489-039A-2803									
Query Match 9.8%; Score 119.4; DB 4; Length 651;									
Best Local Similarity 54.1%; Pred. No. 2e-22;									
Matches 243; Conservative 0; Mismatches 206; Indels 0; Gaps 0;									
Qy	488	GGACAAAAAGCTGTGCTGATGATGATACAGACTCAGTGGTTCAAGGAAAAAGTACTCTT	547						
Db	118	GGCCATCGCGCGTGGTGCTGTGGTTTACCGGCCCTTCGGGGTGGGTAATCCACCGTC	177						
Qy	548	GCATGTGCACCTGAGTCGTGAGTTGCTATGCAAGAGCCACCTCACGTATGTACTTGATGCT	607						
Db	178	GCCGGGGCGCTGGAGGAGGCCCTTGCATGAGCGCGGCTCAGCACTATCTGCTGGACGGC	237						
Qy	608	GACAACCTCAGACATGCGCTTAATAGAGATTTAAGCTTTAAGGCAGAAAGCCGTGCAGAA	667						

Search completed: September 9, 2005, 12:44:43
Job time : 256 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 9, 2005, 09:55:25 ; Search time 830 Seconds
(without alignments)
9629.315 Million cell updates/sec

Title: US-10-829-432-3
Perfect score: 1217
Sequence: 1 ggcgtcggttcattcatca.....aaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US07_PUB_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1217	100.0	1217	US-10-829-432-3	Sequence 3, Appli
2	1037.4	85.2	1291	US-10-425-115-116953	Sequence 116953,
3	859.2	70.6	1335	US-10-425-114-941	Sequence 941, App
4	704.8	57.9	1498	US-10-425-115-116950	Sequence 116950,
5	604.2	49.6	1498	US-10-767-701-9302	Sequence 9302, Ap
6	555	45.6	1319	US-10-739-930-5324	Sequence 5324, Ap
7	387	31.8	548	US-10-437-963-72904	Sequence 72904, A

ALIGNMENTS

RESULT 1

US-10-829-432-3
; Sequence 3, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1217
; TYPE: DNA
; ORGANISM: Zea mays
US-10-829-432-3

Query Match 100.0%; Score 1217; DB 19; Length 1217;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGTCCGTTTCATTTCATCAACAGAACCTCTGTCTACACACGACGACACC 60
|||||

Sequence 9, Appli
Sequence 51779, A
Sequence 5283, Ap
Sequence 62905, A
Sequence 7, Appli
Sequence 62889, A
Sequence 18247, A
Sequence 1150, Ap
Sequence 135127,
Sequence 119620,
Sequence 9652, Ap
Sequence 27400, A
Sequence 1, Appli
Sequence 16540, A
Sequence 710, App
Sequence 119621,
Sequence 869, App
Sequence 65, Appl
Sequence 1680, Ap
Sequence 1680, Ap
Sequence 126988,
Sequence 4843, Ap
Sequence 45866, A
Sequence 29593, A
Sequence 35035, A
Sequence 35036, A
Sequence 5, Appli
Sequence 1323, Ap
Sequence 16007, A
Sequence 62888, A
Sequence 35733, A
Sequence 15460, A
Sequence 41542, A
Sequence 11, Appl
Sequence 41371, A
Sequence 1295, Ap
Sequence 34804, A

QY 603 ATGGTGACAACTCTCAGACATGGCCTAAATAGAGATTTAAAGCTTTAAAGGACAGAACCGGTG 662
DB 542 ATGGTGACAACTCTCAGACATGGCCTAAATAGAGATTTAAAGCTTTAAAGGACAGAACCGGTG 601
QY 663 CAGAAAATATACGAAGAGTTGGTGAAGTGCAGAAAGCTTTTCTGATGCTGCTGTCATAT 722
DB 602 CAGAAAATATACGAAGAGTTGGTGAAGTGCAGAAAGCTTTTCTGATGCTGCTGTCATAT 661
QY 723 GCATTGCTAGTCTGATATCTCCATACAGGAGAGTCTGATGCATGCCGTGCTCTACTTTC 782
DB 662 GCATTGCTAGTCTGATATCTCCATACAGGAGAGTCTGATGCATGCCGTGCTCTACTTTC 721
QY 783 CACATCTCAACTTTATTTGAAGTATTTATTTGATTTGCCCTTAAAAATTTTGTGAAGCTCGTG 842
DB 722 CACATCTCAACTTTATTTGAAGTATTTATTTGATTTGCCCTTAAAAATTTTGTGAAGCTCGTG 781
QY 843 ATCTTAAAGGCTTATACAGCTTGACGCTACAGGAGAGATTTAAAGGTTTCACTGGAATTG 902
DB 782 ATCTTAAAGGCTTATACAGCTTGACGCTACAGGAGAGATTTAAAGGTTTCACTGGAATTG 841
QY 903 ATGATCCATACGAAACCAACCAATTAATGCTGAGATAGTAAATTAAGATGAAAGATGAGGAAT 962
DB 842 ATGATCCATACGAAACCAACCAATTAATGCTGAGATAGTAAATTAAGATGAAAGATGAGGAAT 901
QY 963 GCCCTTACCCAAAGCAATGGCCAAAGTCTATGCTACCTTGAAGAAACCGATATT 1022
DB 902 GCCCTTACCCAAAGCAATGGCCAAAGTCTATGCTACCTTGAAGAAACCGATATT 961
QY 1023 TGCAGCTTAGTATATGATTTTGAAGATGATCTGATCTGATCTGCTGCTGCTACTTACTTG 1082
DB 962 TGCAGCTTAGTATATGATTTTGAAGATGATCTGATCTGATCTGCTGCTGCTACTTACTTG 1021
QY 1083 TGGACACAATAAGATCTGTTGTTGGTCAATCAATGAATAAAGGATCAACATGTAGGAAGTA 1142
DB 1022 TGGACACAATAAGATCTGTTGTTGGTCAATCAATGAATAAAGGATCAACATGTAGGAAGTA 1081
QY 1143 ACAGAGGTGACGGTTCATT--CAGAAACGATATGATTCATTCTGTTAA 1190
DB 1082 ACAGAGGTGACGGTTCATTCTGATCAGCAATTGAGATGATCTGCTGTTAA 1131

RESULT 3

US-10-425-114-941
; Sequence 941, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 941
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700097023_FLI
US-10-425-114-941

Query Match 70.6%; Score 859.2; DB 18; Length 1335;
Best Local Similarity 91.2%; Pred. No. 5e-229;
Matches 979; Conservative 0; Mismatches 13; Indels 82; Gaps 3;
QY 184 CACCCCTAGCGATGGCCACTCCCGCGCTCGTGATCCATGCGCTCACTCCCGTTCCTC 243
DB 1 CACCCCTAGCGATGGCCACTCCCGCGCTCGTGATCCATGCGCTCACTCCCGTTCCTC 60

RESULT 4

US-10-425-115-116950
; Sequence 116950, Application US/10425115

QY 244 ACACCTCTTCCGCGGTCTCGCAGTGATGTGGGCGCCGCGAGGGGAGGGCCGCGGTGC 303
DB 61 ACACCTCTTCCGCGGTCTCGCAGTGATGTGGGCGCCGCGAGGGGAGGGCCGCGGTGC 120
QY 304 CGGTACGACATGCCACCGCGGATTTGGCGGTGGTGGCGCGCGCGCGCGGATGGAGC 363
DB 121 CGGTACGACATGCCACCGCGGATTTGGCGGTGGTGGCGCGCGCGCGGATTTGGAGC 180
QY 364 AGCGCC - CGGGGAGGCGCCGACAGCCAGTGAAGGAGAGCCTGTAATGTGCAACATTTG 422
DB 181 AGCGCCAGGAGGCGCCGACAGCCAGTGAAGGAGAGCCTGTAATGTGCAACATTTG 240
QY 423 GGAATCGACATAATATTTTATGGCACAATTCCTGATTGGACAATCTCGATAGACAGAAAT 482
DB 241 GGAATCGACATAATATTTTATGGCACAATTCCTGATTGGACAATCTCGATAGACAGAAAT 300
QY 483 TGCTGGGCAAAAAGGCTGTGCTGATATGATAACAGGACTCAGTGGTTCAGGGAAGATA 542
DB 301 TGCTGGGCAAAAAGGCTGTGCTGATATGATAACAGGACTCAGTGGTTCAGGGAAGATA 360
QY 543 CTCTTGCAATGCTGCTGAGTCTGATTTGCAATTTGAGAGGCGCCCTCAGCTATGTACTTG 602
DB 361 CTCTTGCAATGCTGCTGAGTCTGATTTGCAATTTGAGAGGCGCCCTCAGCTATGTACTTG 420
QY 603 ATGTTGACAACCTCAGACATGGCCTAAATAGAGATTTTAAAGCTTTTAAAGGACAGAACCGGTG 662
DB 421 ATGTTGACAACCTCAGACATGGCCTAAATAGAGATTTTAAAGCTTTTAAAGGACAGAACCGGTG 480
QY 663 CAGAAAATATACGAAGAGTTGGTGAAGTGCAGAAAGCTTTTGTGCTGATGCTGCTGCTCATAT 722
DB 481 CAGAAAATATACGAAGAGTTGGTGAAGTGCAGAAAGCTTTTGTGCTGATGCTGCTGCTCATAT 540
QY 723 GCATTGCTAGCTTGCATATCTCCATACAGGAGAGATCGTGATGCGTCCGCTCTACTTTC 782
DB 541 GCATTGCTAGCTTGCATATCTCCATACAGGAGAGATCGTGATGCGTCCGCTCTACTTTC 600
QY 783 CACATCTCAACTTTTATTTGAAGTATTTTATTTGATTTGCCCTTAAAAATTTTGTGAAGCTCGTG 842
DB 601 CACATCTCAACTTTTATTTGAAGTATTTTATTTGATTTGCCCTTAAAAATTTTGTGAAGCTCGTG 660
QY 843 ATCTTAAAGGCTTATACAGCTTGCAGCTACAGGAAAGATTTAAAGGTTTCACTGGAATTG 902
DB 661 ATCTTAAAGGCTTGCATCAAGCTTGCAGCTACAGGAAAGATTTAAAGGTTTCACTGGAATTG 720
QY 903 ATGATCCATACGAAACCAACCAATTAATGCTGAGATAAGGATATATATATCATCTGGCA 932
DB 721 ATGATCCATACGAAACCAACCAATTAATGCTGAGATAAGGATATATATATCATCTGGCA 780
QY 933 -----AGATAGTAATT 943
DB 781 TTGTTTTCAAGGATTAACCTTGAAGTCTTCTCCGCTGAGATTTAAAGTAGTAATT 840
QY 944 AAGATGAAGATAGGAAATGCCCTTACCCAAAGCAATGGCCAAAGCAAGTTCTATGCTAC 1003
DB 841 ACGATGAAGATAGGAAATGCCCTTACCCAAAGCAATGGCCAAAGCAAGTTCTATGCTAC 900
QY 1004 CTTGAAGAAACCGATATTTGCAAGCTTAGTATATGATTTTGTGAGAGATGATCTGATT 1063
DB 901 CTTGAAGAGAACGGATATTTGCAAGCTTAGCATATGATTTTGTGAGAGATGATCTGATT 960
QY 1064 CTTGTTGTTCCATTACTTTGCGACACATAAGATCTGTTGTTGGTTCACATGAATAAAGG 1123
DB 961 CT--TGTGTTCCATTACTTTGCGACACATAGCTCTGTTGTTGGTTCACATGAATAAAGG 1018
QY 1124 CATCAACATGTAGGAAGTAAACAGAGGTACCGTTTCATTTCAGAAACCGGATATGGA 1177
DB 1019 CATCAACACGTAGGAAGTAAACAGAGGTACCGTTTCATTTCAGAAACCGGATATAGA 1072

```
; Publication NO. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 116950
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_38149C.1
US-10-425-115-116950

Query Match      57.9%; Score 704.8; DB 20; Length 1498;
Best Local Similarity 82.2%; Pred. No. 7.7e-186;
Matches 953; Conservative 0; Mismatches 22; Indels 185; Gaps 5;

QY 25 AACAGAACCTCTGGTCACACACAGCAGCAACACCGAGCCAGCGCCGCCGCCAGCCAG 84
DB 21 AACAGAACCTCTGGT--CACACACGAGCAAGCAGCGCCAGCGCCGCCGCCAGCCAC 78
QY 85 CCAGGGCCAAACGCAAGGCAACACCTCTCTAGCGCGAGCGCGAGCGCTGCGCGTATCCT 144
DB 79 -----GGCAACACCTCTCTAGCGCGAGCGCGCGCTGCGCGTATCCT 122
QY 145 CGTAATCCACAGCGCGCTCCGCTCC-----TCCAGGCTCCACCCCTAGCGATGC 198
DB 123 CGTAAGTCCACAGCGCGCGCTCCGCTCTCCAGTCCAGGTCTCACCCCTAGCGATGC 182
QY 199 GCCACTCCCGCGCGCTGTCATCATGCGCTCACTCCCGCTTCTTCCACTCTTCCGCGGG 258
DB 183 GCCACTCCCGCGCGCTGTCATCATGCGCTCACTCCCGCTTCTTCCACTCTTCCGCGGG 242
QY 259 TCTCGCAGTATGATGGCGCGCGAGGGAGGCGCGCGTGGCGGTACGCACTGCGCA 318
DB 243 CCTCGCAGCGATAGTGGCGCGCGAGGGAGGCGCGCGTGGCGGTACGCACTGCGCA 302
QY 319 CCGCGCATTTGGCGGTGGTGGTGGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 377
DB 303 CCGCGCATTTGGCGGTGGTGGTGGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 362
QY 378 CCGCGCACGCCAGTGAAGGAGAGCGCTGTAATGTGCAACATTTGGGAAATCGACTAATA 437
DB 363 CCGCGCACGCCAGTGAAGGAGAGCGCTGTAATGTGCAACATTTGGGAAATCGACTAATA 422
QY 438 TTTTATGGCACAATTGCTGTTGATGACCAATCTGATAGACAGAAATGCTGGGACAAAAG 497
DB 423 TTTTATGGCACAATTGCTGTTGATGACCAATCTGATAGACAGAAATGCTGGGACAAAAG 482
QY 498 GCTGTGCTGTATGATGATAACAGACTCAGTGTTTACAGGGAAGTACTCTTGCATGTGCAC 557
DB 483 GCTGTGCTGTATGATGATAACAGACTCAGTGTTTACAGGGAAGTACTCTTGCATGTGCAC 542
QY 558 TGAGTCGTGAGTTGTCATTCAGAGGCCACCTCACGTATGTACTCATGTGTGACCACTCA 617
DB 543 TGAGTCGTGAGTTGTCATTCAGAGGCCACCTCACGTATGTACTCATGTGTGACCACTCA 602
QY 618 GACATGGCCCTAAATAGAGATTTTAAGCTTTAAGCGAGAGACCGTGCAGAAATATACGAA 677
DB 603 GACATGGCCCTAAATAGAGATTTTAAGCTTTAAGCGAGAGACCGTGCAGAAATATACGAA 662
QY 678 GAGTTGGTGAAGTGGCAAGAGCTTTTTCGTGATGCTGGTGTCTATATGCAATTCGTAGTTGA 737
DB 663 GAGTTGGTGAAGTGGCAAGAGCTTTTTCGTGATGCTGGTGTCTATATGCAATTCGTAGTTGA 722
QY 738 TATCTCCATACAGGAGAGATCGTGTATGATGCGGCTGCTTCTTCCACATTCCTAACTTTA 797
```

```
DB 723 TATTGACATACAGGAGAG----- 740
QY 798 TTGAAGTATTATTGATTGGCCCTAAAAATTTGTGAAGCTCGTGATCCTTAAAGCCCTAT 857
DB 741 ----- 740
QY 858 ACAAGCTTGACGTACAGGAAAGATTAAAGGTTTCTACTGGAATTTGATGATCCATACGAAC 917
DB 741 -----AC 742
QY 918 CACCAATTAATGTCGATAGTAATTAAGATGAAGATGAGGAATGCCCTTCAACCCAAAG 977
DB 743 CATCAATTAATGGTGAGATAGTAATTAAGATCTAAGATGGGGAATGCCCTTCAACCAATAG 802
QY 978 CAATGGCCCAAGCAAGTTCTATGCTACCTTGAAGAAACGGATATTTGCAAGCTTAGTATA 1037
DB 803 CAATGGCCCAAGCAAGTTCTATGCTACCTTGAAGAAACGGATATTTGCAAGCTTAGTATA 862
QY 1038 TGTATTTTGAAGATGATCTGATTTCTTGTGTCTCCATTACTTGTGGACACAATAAGAT 1097
DB 863 TGTATTTTGAAGATGATCTGATTTCTTGTGTCTCCATTACTTGTGGACACAATAAGAT 922
QY 1098 CTGTTGTTGGTCAATCAATTAAGAGGATCAACATGTAGGAAGTAACAGAAAGGTACGGTT 1157
DB 923 CTGTTGTTGGTCAATCAATTAAGAGGATCAACATGTAGGAAGTAACAGAAAGGTACGGTT 982
QY 1158 CATTCAAGAAACGGATATGGA 1177
DB 983 CATTCAAGAAACGGATATAGA 1002

RESULT 5
US-10-767-701-9302
; Sequence 9302, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 9302
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-ClUS114933_1
US-10-767-701-9302

Query Match      49.6%; Score 604.2; DB 19; Length 945;
Best Local Similarity 91.6%; Pred. No. 7.7e-158;
Matches 664; Conservative 0; Mismatches 53; Indels 8; Gaps 2;

QY 466 ATCTGATAGACAGAAATTTGCTGGGACAAAAGGCTGTGCTGATGGATAACAGGACTCAG 525
DB 1 ATCTGATAGACAGAAATTTGCTGGGACAAAAGGCTGTGCTGATGGATAACAGGACTCAG 60
QY 526 TGGTTCAGGGAAGTACTCTTGTGTCATGAGTGTGATGCTGAGTGTGATGCTGAGTGTGATGAGGCCA 585
DB 61 TGGTTCAGGGAAGTACTCTTGTGTCATGAGTGTGATGCTGAGTGTGATGAGGCCA 120
QY 586 CCTCAGCTATGTCATGATGTCACCACTCAGACATGGCCCTTAATAGAGATTTAAGCTT 645
DB 121 CATCAGCTATGTCATGATGTCACCACTTAGACATGGCCCTCAATCAGAGATTTAAGCTT 180
QY 646 TAAGCAGAACGCGTGCAGAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCG 705
DB 181 TAAGCAGAACGCGTGCAGAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCG 240
```

```
QY 706 TGATGCTGGTGCATATGCAATGCTAGCTTGTATCTCCATACAGGAGAGATCGTGATGC 765
Db 241 CGATGCTGGTATCATATGCATTTGCTAGCTGTATATCTCCATACAGGAGAGATCGTGATGC 300
QY 766 ATGCCGTGCTTACTTCCACATCTCAATCTTATTTAGTGAATTTATTTGCTCCCTAAA 825
Db 301 ATGCCGTGCTTACTTCCAGATTTCAACTTTATTTAGTGAATTTATTTGCTCCCTAAA 360
QY 826 AATTTTGAAGCTCGTGATCCTAAAGCCTATACAGCTTGCACGTFACAGGAAAGATTAA 885
Db 361 AATTTTGAAGCTCGTGATCCTAAAGCTGTGTACAGCTTGCACGACAGGAAAGATTAA 420
QY 886 AGTTTCACTGGAATTTGATGATCCATACGAACCAACCAATTAATGTTGAGATAGTAATTAA 945
Db 421 AGTTTCACTGGAATTTGATGATCCATACGAACCGCCAGTTAATGTTGAGATAGTAATTAC 480
QY 946 GATGAAGATGAGGATGCCCTTCCACCAAGCAATGGCCAAGCAATGCTTATGCTACCT 1005
Db 481 GATGAAGATGAGGATGCCCTTCCACCAAGCAATGGCCAAGCAATGCTTATGCTACCT 540
QY 1006 TGAAGAAAACGGATATTTGCAAGCTTAGTATATGATTTTGAAGATTTGATCTGATCT 1065
Db 541 TGAAGAAAACGGATATTTGCAAGCTTAGTATATGATTTTGAAGATTTGATCTGATCT 597
QY 1066 TGTGTGCTCATTACTTGTGGACACATAAAGATCTGTTGTTGTCATCATGAATAAAGGCA 1125
Db 598 TGTGTGCTCATTACTTGTGGACACATAAAGATCTGTTGTTGTCATCATGAATAAAGGCA 652
QY 1126 TCAACATGTAGGAAGTAACAGAGGTAACGGTTCATTTCAGAAACGGATATGGAATTCATTGG 1185
Db 653 TCAACATGTAGGAAGTAACAGAGGCGCAGTTTGTTCAGAAACGGATATGGAATTCATTGG 712
QY 1186 TTTAA 1190
Db 713 TTTC A 717

RESULT 6
US-10-739-930-5324
; Sequence 5324, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; NUMBER OF SEQ ID NOS: 2003-12-18
; SEQ ID NO 5324
; LENGTH: 1319
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER7564_1
US-10-739-930-5324

Query Match 45.6%; Score 555; DB 20; Length 1319;
Best Local Similarity 84.7%; Pred. No. 5.3e-144;
Matches 636; Conservative 0; Mismatches 110; Indels 5; Gaps 1;

QY 378 CCCCGCACAGCCAGTGAAGGAGAGCCCTGTAATGTCGAACATTCGGAAATCGACTAATA 437
Db 155 CACTCGCGCCGACTGCGAGAGAGCCCTGTCATGTCGAACATTCGGAAATCGACTAATA 214
QY 438 TTTTATGGCAATTCCTTGAATGGAACAATCTGTATAGACAGAAATTCCTGGGACAAAAG 497
Db 215 TTTTATGGCAATTCCTTGAATGGAACAATCTGTATAGACAGAAATTCCTGGGACAAAAG 274
QY 498 GCTGTGCTGTATGATTAACAGACACTCAGTGGTTCAGGGAAGATCTTGTGATGTGCAC 557
Db 275 GATGTGTCATATGGATACAGGACTCAGCGGTTTCAGGGAAGATCTTGTGATGTGCAC 334
```

```
QY 558 TGAGTCGTGAGTTGCATTTGCAGAGCCACCTCACGTATGTACTTGTGATGGTGACAACCTCA 617
Db 335 TGAGTCGGGAATTTACACTACAGAGCCACACAGTATGTCTTGTGATGGTGACAACCTCA 394
QY 618 GACATGGGCTTAAATAGAGATTTAAGCTTTAAGGCGAAGAGACCGTGCAGAAAATATACGAA 677
Db 335 GACATGGGCTTCAATCGAGATCTAAGCTTTCAAGGCGAAGAGACCGTGCAGAAAATATACGAA 454
QY 678 GAGTTGGTGAAGTGGCAAGCTTTTGTGCTGATGCTGGTGTGCATATGCATGCTAGCTTGA 737
Db 455 GAGTTGGTGAAGTGGCAAGCTTTTGTGCTGATGCTGGTGTGCATATGCATGCTAGCTTGA 514
QY 738 TATCTCCATACAGGAGAGATCGTGATGATGCTGCGCTGCTTACTTCCACATTTCTAACTTTA 797
Db 515 TATCTCCATACAGGAGAGACCGTGTGATGATGCTGCGAGCTTACTTCCAGATTTCTAGATT 574
QY 798 TTGAAGTATTTTGTGATTTGGCCCCATAAAATTTGTGAAGCTCGTGATCTCTAAAGGCTTAT 857
Db 575 TTGAAGTATTTTGTGATTTGGCACTAGAAATTTATGTGAAGCTCGTGATCTCTAAAGGCTTAT 634
QY 858 ACAAGCTTGCACGTACAGGAAAGATTAAAGTTTTCACCTGGAATTCATGATCCATACGAAC 917
Db 635 ACAAGCTTGCACGTACAGGAAAGATTAAAGGTTTTCACCTGGAATTCATGATCCATACGAAT 694
QY 918 CACCAATTAATGCTGAGATAGTAAATTAAGATGAAGATGAGGAATGCCCTTCCACCAAG 977
Db 695 CACCAATTAATGCTGAGATAGTAAATTAAGATGAGGAATGCCCTTCCACCAAG 754
QY 978 CAATGGCCAAAGCAATGCTTCTGCTACCTTGAAGAAACGGATATTTGCAAGCTTAGTATA 1037
Db 755 CAATGGCCAAAGCAATGCTTCTGCTACCTTGAAGAAATGGAATTTGCAAGCTTAGTATA 814
QY 1038 TGTATT-----TTGAGAGATTTGATCTGATCTTCTGCTGCTCATTTGTTGGACACAAT 1092
Db 815 CATACTCCAGATCCAGAAAGATTGATCTTATTTCTTCTGCTGCTCATTTGTTGGACACAAG 874
QY 1093 AAGATCTGTTGTTGTCACATGAATAAAGG 1123
Db 875 TACAATCAATTTGTTCCATCCGGAATAAAG 905

RESULT 7
US-10-437-963-72904
; Sequence 72904, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 72904
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73237C.1
US-10-437-963-72904

Query Match 31.8%; Score 387; DB 19; Length 548;
Best Local Similarity 81.7%; Pred. No. 2.9e-97;
Matches 447; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
```



```
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 51779
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5413C.1
US-10-437-963-51779

Query Match      28.9%; Score 352.2; DB 19; Length 915;
Best Local Similarity 69.6%; Pred. No. 2.1e-87;
Matches 492; Conservative 0; Mismatches 213; Indels 2; Gaps 1;

QY 329 GGGCGGTGGTGGCGGGCGGGGGAATGAGCAGCCCGGGGAGCCCGCCGACAG- 387
DB 209 GGTGAGGCGCGGTGCGCGGGCGGGCCCGGTGACGGGAAGGACGACGATGTGT 268
QY 388 -CCCACTGAAGGAGAGCCCTGTAATGTCGAACATTTGGGAAATCGACTAATATTTATGGC 446
DB 269 TCCAGTTTTTAAACAAGCTCATGCTCAACTGTCGTTAAATCAACAACATCCGATGGC 328
QY 447 ACAATGCTTGATTTGGACAATCTGATAGACAGAAATGCTGGGACAAAAGGCTGTGCG 506
DB 329 ATGACTGCCCACTAAACACAGCTTGACAGACAAAAGTTGCTGAACCCAGAAAGGCTGTGTTG 388
QY 507 TATGATACAGACTCATGCTGTTTCAAGGAAAGTACTTTCATGTCGACTGATGCTGTG 566
DB 389 TTTGATCACTGGCTTAAGTGTGTTGAGGAAAGACACCCCTTGCATGTGCACTGAGCCGTG 448
QY 567 AGTTGATTCAGAGAGCCACCTCACGTATGATCTTGTGTCACAACTTCAGACATGACC 626
DB 449 AGCTGCACTCAAGAGGCACTGACCTATGTTCTTGACGGCGACAAATCTCCGGACAGGCC 508
QY 627 TAAATAGAGATTTAAGCTTTAAGGCGAGAGACCGTGCAGAAATATACGAAGAGTTGGTG 686
DB 509 TGAACAAAGATCTCAGCTTCAAGGCAAGGATCGTGCCGAAATATATCGCAGAGTTGGAG 568
QY 687 AGTGGCAAGCTTTTGTGATGCTGTTGATGATGATGATGATGATGATGATGATGATGAT 746
DB 569 AAGTGGCAAGCTTTTGTGATGCTGTTGATGATGATGATGATGATGATGATGATGATGAT 628
QY 747 ACAGGAGAGATCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 806
DB 629 ATAAGAGTATGAGAGCGCTGCGCAAAATTAATACCAAAATCTTCGTTTCAATGAGGTGT 688
QY 807 TTAATGATTTGCCCTTAAATAATTTGTGAAGCTCGTGATCCTTAAAGGCTTATACAGCTTG 866
DB 689 TCCTGAATGTCCTCACTTGAAGTATGTGAAGAAAGGATCCAAAGGCTGTACAAAGCTTG 748
QY 867 CAGGTACAGGAAGATTAAGTTTCACTGGAATTTGATGATGATGATGATGATGATGATGAT 926
DB 749 CTCGTGCGGCAAAATCAAGGCTTTACGGGAATAGATGATGATGATGATGATGATGATGAT 808
QY 927 ATGTTGAGATAGTAAATTAAGATGAAGATGAGGAATGCGCTTCAACCAAGCAATGGCCA 986
DB 809 ATTGTGAGATTTGATACAGTGCAGAAAGTTGGGACTGCCCTTCACTTAATCAATGGCTG 868
QY 987 AGCAAGTTCTATGCTACCTTGAAGAAAACGGATATTTGCAAGCTTTAG 1033
DB 869 ATCAAGTAGTGCATATCTTGAAGCAATGGATCTTTTCAGAACTAG 915
```

RESULT 10

```
US-10-260-238-5283/c
; Sequence 5283, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
```

RESULT 11

```
US-10-424-599-62905
; Sequence 62905, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 62905
```

```
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Rieke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 5283
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Zea mays
US-10-260-238-5283
```

```
Query Match      28.8%; Score 350.2; DB 17; Length 378;
Best Local Similarity 98.7%; Pred. No. 4.5e-87;
Matches 374; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 245 CACTTTCGCGGGTCTGCCAGTGATAGTGGCGCCGCGAGGGGAGGGCGCGGTGCG 304
DB 378 CACTTTCGCGGGTCTGCCAGCGATAGTGGCGCCGCGAGGGGAGGGCGCGGTGCG 319
QY 305 CGTAGGCACTCCACCCGGGCAATTTGGCGGTGGTGGCGCGCGCGCGCGGGAATGGAGCA 364
DB 318 CGTAGGCACTCCACCCGGGCAATTTGGCGGTGGTGGCGCGCGCGCGCGGGAATGGAGCA 259
QY 365 GCGCC-CGGGAGAGCCCGCAGCCAGCCAGTGAAGAGAGAGCTGTAAATGTCGAACATTGG 423
DB 258 GCGCCAGGGGATGCGCCGCGCAGCCAGCCAGTGAAGAGAGAGCTGTAAATGTCGAACATTGG 199
QY 424 GAAATCGACTAATATTTTATGGCAAAATTTGCTGATGGCAAAATCTGTATAGACAGAAAT 483
DB 198 GAAATCGACTAATATTTTATGGCAAAATTTGCTGATGGCAAAATCTGTATAGACAGAAAT 139
QY 484 GCTGGGCAAAAAGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 543
DB 138 GCTGGGCAAAAAGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 79
QY 544 TCTTGCATGTGCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 603
DB 78 TCTTGCATGTGCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 20
QY 604 TGGTGACAACCTCAGACAT 622
DB 19 TGGTGACAACCTCAGACAT 1
```

```
; LENGTH: 1256
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_27817C.1
US-10-424-599-62905

Query Match      28.1%; Score 342; DB 18; Length 1256;
Best Local Similarity 71.9%; Pred. No. 1.8e-84;
Matches 447; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 410 ATGTCGAACATTGGGAATCGCAATATATTTATGGCACAATTTGCTTGATGGACAATCT 469
Db 200 ATGTCAAATGTTGGCAACTCGACAATAATATATGTCGATGACTGTCCAAATTCAGAAACAA 259

QY 470 GATAGACAGAATTTCTGGGACAAAAGGCTGTGCTGATGATGATAACAGGACTCAGTGGT 529
Db 260 GATAGACAGCAGCTGCTTCAGCAAAAAGGCTGTGTTATATGGCTAACTGGGCTCAGTGGT 319

QY 530 TCAGGGAAAAGTACTCTTTCGATGTGCACCTGAGTCTGAGTTGCATTGCGAGAGGCCACCTC 589
Db 320 TCAGGAAAAGACACTATTTCGATGTGCTTTGAGTCGAAGCTTGCACTCCAAAGGAAAACCTG 379

QY 590 ACGTATGACTTGTGATGGTGACAACTCAGACATGCGCTAAATPAGAGATTTTAAGCTTTAAG 649
Db 380 TCTTACATCTTGTATGGTGACAAATATTCGGCATGCTTAAACCAAGATCTTTAGTTTCAGA 439

QY 650 GCAGAGACCGTGCAGAAAATATACGAAGAGTTGCTGGAAGTGGCGCAAAAGCTTTTTCCTGAT 709
Db 440 GCAGAGAGATCGTTCTGAAAACATTAGAAGGATTTGCTGAGGTGGCGCAAAAGCTTTTTCAGAT 499

QY 710 GCTGCTGTGCATATGCATTGCTAGCTTGATATCTCCATACAGAGAGATCGTGATGCATGC 769
Db 500 GCTGCTGTTATTTGTCATCACTAGTTTATATACCATACCAAAAGATAGAGATGCATGC 559

QY 770 CGTCTCTACTTCCACATCTTAACCTTTATTTGAAGTATTTATTTGATTTGCCCTTAAAGAT 829
Db 560 AGAGCACTAAATTCCAAAGGAGATTTTATTTAGAGTTTTCATAGATGTTTCCACTACATGTG 619

QY 830 TGTGAAGCTCGTATCCTTAAGGCTTATACAGCTTGCACGTACAGGAAAGATTAAGGT 889
Db 620 TGTGAAGCTAGGGACCCAAAGGACTCTACAAGCTTTGCTCGAGCTGGAAAGATCAAAAGT 679

QY 890 TTCCTGGAATTTGATGATCCATACGAAACCAACCAATTAATTTGATGATAGTAAATTAAGATG 949
Db 680 TTCCTGGAATTTGATGATCCATATGACCAACCGTGTAGTTGTGATAGTATTACACAG 739

QY 950 AAAGATAGGAATGCGCTTCCACCAAGCAATGGCCAAAGCAAGTTCTATGCTACCTTGA 1009
Db 740 AAAGGAAGTACTGTAAAGTCTCCAGTGATATGGCTGAAGAAGTGATATCTTACTTTGGAG 799

QY 1010 GAACACGGATATTTGCAAGCTT 1031
Db 800 GAGAACGGATACCTGCGGGCTT 821
```

```
RESULT 12
US-10-829-432-7
; Sequence 7, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 14
```

```
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Glycine max
US-10-829-432-7

Query Match      27.8%; Score 338.8; DB 19; Length 936;
Best Local Similarity 71.5%; Pred. No. 1.2e-83;
Matches 445; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 410 ATGTCGAACATTGGGAATCGCAATATATTTATGGCACAATTTGCTTGATGGACAATCT 469
Db 313 ATGTCAAATGTTGGAACTCGACAATAATATATGTCGATGACTGTCCAAATTCAGAAACAA 372

QY 470 GATAGACAGAATTTGCTGGGACAAAAGGCTGTGCTGATGATGATAACAGGACTCAGTGGT 529
Db 373 GATAGACAGCAGCTGCTTCAGCAAAAAGGCTGTGTTATATGGCTAACTGGGCTCAGCGGA 432

QY 530 TCAGGGAAAAGTACTCTTTCGATGTGCACCTGAGTCTGAGTTGCATTGCGAGAGGCCACCTC 589
Db 433 TCAGGAAAAGACACTATTTCGATGTGCTCTGAGTCAAAGCTTGCACTCCAAAGGAAAACCTG 492

QY 590 ACGTATGACTTGTGATGGTGACAACTCAGACATGCGCTTAAATPAGAGATTTTAAGCTTTAAG 649
Db 493 TCTTACATCTTGTATGGTGACAAATATTCGGCATGCTTAAACCAAGGATCTTTAGTTTTCAGA 552

QY 650 GCAGAGACCGTGCAGAAAATATACGAAGAGTTGCTGGAAGTGGCGCAAAAGCTTTTTCCTGAT 709
Db 553 GCAGAGAGATCGTTCTGAAAACATTAGAAGGATTTGCTGAGGTGGCGCAAAAGCTTTTTCAGAT 612

QY 710 GCTGCTGTGCATATGCATTGCTAGCTTGATATCTCCATACAGAGAGATCGTGATGCATGC 769
Db 613 GCTGCTGTTATTTGTCATCACTAGTTTATATACCATACCAAAAGATAGAGATGCATGC 672

QY 770 CGTCTCTACTTCCACATCTTAACCTTTATTTGAAGTATTTATTTGATTTGCCCTTAAAGAT 829
Db 673 AGAGCACTACTTTCAAAGGAGATTTTATTTAGAGTTTTCATAGATGTTTCCACTACATGTG 732

QY 830 TGTGAAGCTCGTATCCTTAAGGCTTATACAGCTTGCACGTACAGGAAAGATTAAGGT 889
Db 733 TGTGAAGCTAGGGACCCAAAGGACTCTACAAGCTTTGCTCGAGCTGGAAAGATCAAAAGT 792

QY 890 TTCCTGGAATTTGATGATCCATACGAAACCAACCAATTAATTTGATGATAGTAAATTAAGATG 949
Db 793 TTCCTGGAATTTGATGATCCATATGACCAACCGTGTAGTTGTGATAGTATTACACAG 852

QY 950 AAAGATAGGAATGCGCTTCCACCAAGCAATGGCCAAAGCAAGTTCTATGCTACCTTGA 1009
Db 853 AAAGGAAGTACTGTAAAGTCTCCAGTGATATGGCTGAAGAAGTGATATCTTACTTTGGAG 912

QY 1010 GAACACGGATATTTGCAAGCTT 1031
Db 913 GAGAACGGATACCTGCGGGCTT 934
```

```
RESULT 13
US-10-425-115-62889
; Sequence 62889, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 62889
; LENGTH: 2611
```



```
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2611)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_157357C.1
US-10-425-115-62889

Query Match      27.6%; Score 336; DB 20; Length 2611;
Best Local Similarity 63.8%; Pred. No. 1.3e-82;
Matches 528; Conservative 0; Mismatches 295; Indels 5; Gaps 1;

QY 206 CCGGCGCTGTCATCATGGCTCACTCCCGTTCCTCACACTCTTCGCGGGTCTCGCC 265
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1344 CCGGAACCGGAGCCCGCGCGAATCTAGGTTGCTCGCTCGCCCTCTCCCGAG 1403
QY 266 AGTGATAGTGGCGCGCGAGGGGAGGCGCGGTGCGGTACGCACTGCCACCGCGC 325
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1404 ACTCTAGCGGAGGAATTAGGACGAGGTAAACGCGACAGAGGCTGCGAAGGAGCGCC 1463
QY 326 ATTGGCGGTGGTGGCGGCGCGGCGGGAATGAGAGCGCCCGGGAGGCGCCGAC 385
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1464 CCTCGTGAATGCGCGCGCAGCCGATCCGTGGAGGAGCAGC-----CGGAGCAGCAGGA 1518
QY 386 AGCCGAGTGAAGGAGAGCGCTGTAATGTCGAACATTGGGMAATCGACTAATATTTATGG 445
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1519 GTTGAAGGCGAAGCTTTGAAATGTCATCCACTGTGCGGAGTCAATATCTCTTGG 1578
QY 446 CACAAATGCTGATTGGCAATCTGATAGACAGAAATTTGCTGGGACAAAAAGGCTGTGTC 505
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1579 CATGATTGCTAGTTGGCAAGACTGATCGCAGAGTACTCAACAGAGGCTGTGT 1638
QY 506 GTATGGATAACAGGACTCAGTGGTTCAGGGAAGTACTCTTTGCAATGTCACAGTGT 565
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1639 GTCTGGATCACAGGCTTAGTGGCTCAGGTAAAGTACTCTTGGCATGTACATTTAGGCGG 1698
QY 566 GAGTTGCATTGACAGAGGCACTCAGCTATGTACTTGTGTTGATGTTGACAACTCAGACATGC 625
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1699 GAGCTCCACACAGAGGGAAGCTTGCAATGTTCTTGACGGTGAATCTTAAGACATGGT 1758
QY 626 CTAATAGAGATTAAAGCTTTAAAGCAGAAAGACCGTGCAGAAAAATATACGAAGAGTTGGT 685
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1759 CTGAACAGGATCTTGGCTTTAAAGCTGAAGCCGTGCTGAAATATTAAGGAGTGGT 1818
QY 686 GAAGTGGCAAGCTTTTGTGATGCTGGTGTATATGATGCTTGTAGTGTATATCTCCA 745
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1819 GAAGTAGCAAGATTATTTGCAGATCGGCGCTTGTATGTTATGCAAGTTTGATATCTCCA 1878
QY 746 TACAGGAGATCGTGATGATGCGCTGCTACTTCCACATTTCTAACTTTATTTGAAGTA 805
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1879 TATAGGAGACCGTGAATCTTGGCGTGCATTGTTGTCTGATGATGCTTCAATGAAGTT 1938
QY 806 TTTATTGATTGCCCCCTAAATAATTTGTAAGCTCGTGTATCTTAAAGGCTTATACAAGCTT 865
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1939 TTTCTGAACATGCTCTTGGATTATGTAGCAAGAGATCCAAAGGCTCTATAGGCTT 1998
QY 866 GCACGTACAGGAAGATTAAAGGTTTCACTGGAAATGATGATCCATACGAACCACTAAT 925
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1999 GCTCGTCAGGAAAAATAAAGGTTTTACAGGAATAGATGACCTTTATGAAGCACCCTG 2058
QY 926 AATGTTGATAGTAAATTAAGATGAAGATGAGGAATGCCCTTACCCTAAAGCAATGCC 985
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2059 AATTTGATGATGATCAAGAGAGTAGATGTTGTTATGCCCTCCGCTGCTGAGATGCA 2118
QY 986 AAGCAAGTTCTATGCTACTCTTGAAGAAACGATATTTGCAAGCTTAG 1033
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2119 GGGCAAGTTGTTACTTACCTTGAGGAAAGGCTTCTGACAGTAG 2166
```

RESULT 14

US-10-425-114-18247

```
; Sequence 18247, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18247
; LENGTH: 1180
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3079-005-C3_FLI
US-10-425-114-18247

Query Match      27.4%; Score 333; DB 18; Length 1180;
Best Local Similarity 70.2%; Pred. No. 5.6e-82;
Matches 447; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 424 GAAATCGACTAATATTTTATGGCAATTTGCTTGGACAAATCTGATGGACAAATCTGATAGACAAATTT 483
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
328 GAAGTCATCAATATCTTTCGGCATGATTTGCTAGTTGGCAAGACTGATGCCAAGACT 387
QY 484 GCTGGGACAAAAGCGCTGTGTATGATTAACAGGACTCAGTGGTTCAGGGAAAAGTAC 543
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
388 ACTCAACAGAAAGCGTGTGTCTGGATCACAGGCTTAGTGGCTCAGGTAAAAGTAC 447
QY 544 TCTTGCATGTGCATCGTGTGATTTGCAATTCAGAGGCCACCTCACGTATGTACTTGA 603
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
448 CTTGGCATGTACATTAGCGGGAGCTCCACACAAGAGGGAAGCTTGATATTTCTTGA 507
QY 604 TGGTGACAACTCAGACATGSCCTTAATAGAGATTTAAGCTTTAAGGAGAGACCGTGC 663
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
508 CGGTGATAACTTAAGACATGCTGTAACAAGACTTTGGCTTTAAGCTGAAGATCTGTGC 567
QY 664 AGAAAATATACGAAGAGTTGGTGAAGTGGCAAAAGCTTTTCTGATGCTGGTGTATG 723
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
568 TGAATAATTAAGGAGAGTTGGTGAAGTAGCAAAAGTTATTTTCAGACGCTGGCCTGTATG 627
QY 724 CATTGCTAGCTTGATATCTCCATACAGGAGAGATCGTGATCGATCGCTGCTCTACTTCC 783
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
628 TATTGCAAGTTTGATATCTCCATATAGGAGAGCCGTGAATCTTCCCGTGCACCTGTTGTC 687
QY 784 ACATTCTAACTTTATGAAATTTATTTGATTTGCCCTTAAAGTTTCTGGAATTTGA 843
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
688 GGATGGTAGCTTTTATGAAAGTTTCTTGAATATGCTCTTGAATTTATGTGAAGCAAGGA 747
QY 844 TCCTTAAAGGCTTATACAGCTTGCAAGTACAGGAAAGATTAAAGCTTTTCTGGAATTTGA 903
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
748 CCAAGAGGCTCTATAGCTTGTCTGCTGGGAAAATAAAGGTTTTTACAGGAATAGA 807
QY 904 TGATTCATACCAACCACTAATTAATGTTGATAGTATTAATTAAGATGAAGATGAGGAATG 963
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
808 TGACCTTATGAAGCAACCACTGAATTTGTAGATTGAGATCAAGGAAATAGATGGTGTATG 867
QY 964 CCTTTCACCCAAAGCAATGGCCAAAGAGTTCTATGCTACCTTGAAGAAAACGGATATTT 1023
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
868 CCTTTCGCTTCTGACATGGCGGCAAGTGGTTACTTTACCTTGGAGGAAAGGCTTCTCT 927
QY 1024 GCAAGCTTAGTATATGTTTGTGAGAGATTGATCTG 1060
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
928 GCACGAGTAGGCTGACCGCTGAATCGAAGCTACATCCG 964
```

RESULT 15

US-10-425-114-1150
; Sequence 1150, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 1150
; LENGTH: 1312
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700102592_FLI
US-10-425-114-1150

Query Match 27.4%; Score 333; DB 18; Length 1312;
Best Local Similarity 63.4%; Pred. No. 5.9e-82;
Matches 528; Conservative 0; Mismatches 300; Indels 5; Gaps 1;

QY 206 CCGCGCTCGTGATCCATGGCCCTCATCCCGCTTCTCCTACACTCTTCGCGGGGTCTCGCC 265
DB |||||
QY 212 CCGGAACCGGAGCCCGCGGAATCGAGGGTTGCTCGCCCTCCCGCGAG 271
DB |||||
QY 266 AGTGATAGTGGCGCGCGGAGGGCGCGCGTGGCGGTACGACTGCCACCGCGGC 325
DB |||||
QY 272 ACTCTAGCGGAGGAATTAGGACGAGGGTAACAGCGACAGAGGCTGCGAAGGAGCGCC 331
QY 326 ATTGGCGGTGGGTGGCGCGCGGGAATGGAGCAGCGCCCGGGAGGCCCGGCAC 385
DB |||||
QY 332 CTTCGTAATCGCGCGCGCCGATCGTGGAGGAGCAGCGGAGGAGGTTA 391
QY 386 AGCCAGTGAAGGAGGAGCTGTAA-----TGTGAAATTTGGGAAATCGACTAATATTT 440
DB |||||
QY 392 GTCTAGTTGAAGGGAAGCTTTGAAATGTCTATCCACTGTGCGGAAGTCATCAATATCT 451
QY 441 TATGCACAAATTGCTTGTATGGACAACTCTGATAGACAGAAATTTGCTGGGACAAAAGGCT 500
DB |||||
QY 452 TCTGGCATGATTCCTTAGTTGGCAAGACTGATCGCGAGAGCTACTCAACGAGAAAGGCT 511
QY 501 GTGTCGTATGGATAACAGGACTCAGTGGTTACAGGAAAAGTACTCTTGCATGTGCACTGA 560
DB |||||
QY 512 GTGTTGTCTGGATCACAGSCCTTAGTGGCTCAGGTAAAGTACCTTTGGCATGTACATTAG 571
QY 561 GTGCGTGTGATTTGCAGAGSCCACCCTCAGTATGTACTTGTGATGGTGACACCTCAGAC 620
DB |||||
QY 572 GCCGGGAGCTCCACAAAGAGGGAAGCTTGCATATATTTCTTGACGGTGTAACTTTAAGAC 631
QY 621 ATGGCTTAATAGAGATTTAAGCTTTAAGGCAGAGAGCCGTGCAGAAAATATACGAAGAG 680
DB |||||
QY 632 ATGGTCTGNAACAGGATCTTGGCTTTAAGCTGAAAGACCGTGCTGAAAATATAGGAGAG 691
QY 681 TTGGTGAAGTGGCAAGCTTTTGTGTGATGTGGTGTGCATATGCATTTGCTAGCTTGATAT 740
DB |||||
QY 692 TTGGTGAAGTAGCAAAAGTTATTTTCAGACCGGGCTTGTATGTATTTGCAAGTTTGATAT 751
QY 741 CTCCATACAGGAGAGATCGTGATGATGCGCGTCTACTTCCACATTTCTAATTTATTG 800
DB |||||
QY 752 CTCCATATAGGAGAGACCGTGAATCTTGGCGTGCATTTGTATCGGATGGTAGCTTTATG 811
QY 801 AAGTATTTTATGATTTGCCCCATAAAATTTGTGAAGCTCGTGATCTCTAAAGGCCATTACA 860
DB |||||
QY 812 AAGTTTTCTTGATATATGCTTGGAAATTATGTGAAGCAAGGACCCCAAGGGCCTCTATA 871
QY 861 AGCTTGCAGGTACAGGAAGATTAAAGGTTTCACTGGAAATTGATGATCCATACGAACCAC 920

Db |||||
QY 872 AGCTTGTCTGCTGGGAAAAATAAAGGTTTTTACAGGAATAGATGACCCCTTATGAAGCAC 931
QY 921 CAATTATGGTGAGATAGTAAATTAAGATGAAGATGAGGAATGCCCTTCAACCAAGCAA 980
DB |||||
QY 932 CACTGAATTGTGAGATTGAGATCAAGGAAATAGATGGTGTATGGCTTTCGGCTTCTGACA 991
QY 981 TGGCCAAAGCAAGTTCTATGTCTACCTTGAAGAAAACGGGATATTTGCAAGCTTAG 1033
DB |||||
QY 992 TGGGGGGCAAGTGGTTACTTACCTTGAGGAGAAAGGCTTCTGCACGAGTAG 1044

Search completed: September 9, 2005, 12:58:41
Job time : 834 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 9, 2005, 09:31:45 ; Search time 3856 Seconds
(without alignments)
12013.541 Million cell updates/sec

Title: US-10-829-432-3
Perfect score: 1217
Sequence: 1 gcgtcgcttctattcatca.....aaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	701.6	57.6	720	7	COS28976
2	590	48.5	675	2	BE357876
3	583.4	47.9	691	2	BE355111
4	495.8	40.7	1075	7	CK205911
5	486.6	40.0	617	6	CD863007
6	477.2	39.2	645	7	COS32501
7	468	38.5	614	6	CA204452
8	455.4	37.4	799	7	CNI25017
9	446.8	36.7	809	6	CA766480
10	442.2	36.3	860	7	CR285854
11	422.2	34.7	746	2	BE361884
12	415	34.1	673	6	CA182186
13	411.2	33.8	525	6	CA502007
14	409.2	33.6	721	2	AW922946
15	400.2	32.9	699	2	BE361874
16	395.4	32.5	1071	7	CK212359
17	389	32.0	673	6	CD924390
18	381	31.3	496	6	CA497171
19	371.6	30.5	866	9	CG159926
20	371.6	30.5	922	9	CG454522
21	369	30.3	609	9	CG801367
22	356.2	29.3	671	7	CF633032
23	355.4	29.2	708	6	CA762202
24	352.2	28.9	915	9	CL981208

C	25	347	28.5	728	6	CD864184	CD864184 AZO1.109F
	26	344.8	28.3	421	2	BF484142	BF484142 WHE1788 G
	27	336.4	27.6	867	7	CO095086	CO095086 GR_Eal71
	28	334.4	27.5	487	6	CA720307	CA720307 wkmzn.pk0
C	29	333	27.4	629	7	COS32500	COS32500 3530.1.21
	30	332.2	27.3	493	6	CD864183	CD864183 AZO1.109F
C	31	330.6	27.2	627	2	AW052991	AW052991 614077E08
C	32	323.6	26.6	720	7	CK827430	CK827430 zmrsub1.0
C	33	318.8	26.2	591	4	BG873962	BG873962 MEST44-F0
C	34	315.8	25.9	774	4	BM817061	BM817061 HC01C03.T
C	35	311.8	25.6	641	6	CD927759	CD927759 GR45.103B
C	36	311.2	25.6	754	6	CA921172	CA921172 EST638890
	37	310.2	25.5	797	7	CNI29765	CNI29765 RHOH1.37
	38	309.8	25.5	803	7	CO476247	CO476247 QSO0611.B
C	39	309.2	25.4	902	7	CV245034	CV245034 WS0256.B2
	40	309	25.4	696	6	CA085472	CA085472 SCJLAM209
	41	308.2	25.3	837	7	CO200697	CO200697 RTCNT2.1
	42	307.6	25.3	970	3	AY109629	AY109629 Zea mays
C	43	307	25.2	957	7	CV260804	CV260804 WS02015.B
	44	304.6	25.0	639	7	CO529312	CO529312 3530.1.19
	45	304.4	25.0	536	2	AW560397	AW560397 EST315445

ALIGNMENTS

RESULT 1
COS28976
LOCUS
DEFINITION
3530.1.191.1.D07.Y.1.3530 - Full length cDNA library created by
Invitrogen from multiple tissues Zea mays cDNA, mRNA sequence.
ACCESSION
COS28976
VERSION
COS28976.1
KEYWORDS
EST.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 720)
Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 3530.1.191.1 row: D column: 07.
Location/Qualifiers
1. 720
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="B73"
/db_xref="taxon:4577"
/tissue_type="multiple"
/dev_stage="varies by tissue"
/lab_host="DH10B"
/clone_lib="3530 - Full length cDNA library created by
Invitrogen from multiple tissues"
/note="Organ: silks, husks, ears, pollen, shoot tips,
leaf, root tips, whole seed, embryo; Vector: pCMV-SPORT
6.1; Site 1: EcoRV; Site 2: NotI; Maize Gene Discovery
Project contracted with Invitrogen to produce a
normalized, full length library in a pSPORT vector. This
is a Gateway compatible vector, permitting clone movement
to new vector backbones for expression in diverse host
cells using recombination rather than restriction enzymes.
Details of the vector and sequencing primers are available
at ZmDB in the EST library description tables. poly(A) +

FEATURES
source

mRNA was prepared by Invitrogen, and equimolar amounts of RNA from each of the 12 tissue samples were mixed together for selection of mRNA with a 5' cap. After synthesis of cDNA, a normalization step was conducted against the mixture of RNA sources. This step effected a 20X to 80X reduction in common transcript types. Tissues prepared: 1. just emerging silks; 2. inner husks from ears of sample #1; 3. 20 dap aleurone; 4. immature tassels, stages from 1-2 mm to 1-2 cm; 5. 2 mm to 2 cm ears; 6. pollen; 7. 1 cm vegetative shoot tips from 15 day old seedlings; all leaves with an expanded or partially expanded sheath were removed; 8. mature leaf tissue; 9. 0.5 cm long root tips from 15 day old seedlings; 10. 10 dap whole seed; 11. 12 dap endosperm and embryo; 12. 17 dap endosperm and embryo. All of the sequenced clones in project 3530 will be archived at the University of Arizona along with the Unigene clones from the Maize Gene Discovery EST sequencing projects. Clones can be ordered through the ZmDB web site or directly from the University of Arizona (<http://www.genome.arizona.edu/orders/>). High density filters containing over 18,000 clones can also be ordered from the University of Arizona."

ORIGIN

Query Match 57.6%; Score 701.6; DB 7; Length 720;
Best Local Similarity 99.3%; Pred. No. 2.3e-165;
Matches 715; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 82 CAGCAGGGCCAAACGGCAGGCAACACCTCTCTAGCCGAGCGCGCTGCCTCAT 141
DB 1 CAGCCAGGGCCAAACGGCAGGCAACACCTCTCTAGCCGAGCGCGCTGCCTCAT 60

QY 142 CCTCGTAAATCCACAGCGCGCCCTCCGTCCTCCAGGCTCCACCTTAGCGATGCC 201
DB 61 CTTCTGTAATCCACAGCGCGCCCTCCGTCCTCCAGGCTCCACCTTAGCGATGCC 120

QY 202 ACTCCGCGGCTCGTGATCCATGGCTCACTCCCGCTTCTCACACTCTTCGCGGGTCT 261
DB 121 ACTCCGCGGCTCGTGATCCATGGCTCACTCCCGCTTCTCACACTCTTCGCGGGTCT 180

QY 262 CGCCAGTGTAGTGGGCGCGGAGGGGAGGGCGCGGTGGCTAGCGACTGCCACGG 321
DB 181 CGCCAGCGATAGTGGGCGCGGAGGGGAGGGCGCGGTGGCTAGCGACTGCCACGG 240

QY 322 CGGCATTGGCGGCTGGGTGGCGCGGCGGGAATGAGCAGCGCC -CGGGAGGCGCC 380
DB 241 CGGCATTGGCGGCTGGGTGGCGCGGCGGGAATGAGCAGCGCCCGGGATGCC 300

QY 381 CGCACAGCCCACTGAAGGAGAGCCTGTAAATGTCGAACATTGGGAAATCGACTAATTT 440
DB 301 CGCACAGCCCACTGAAGGAGAGCCTGTAAATGTCGAACATTGGGAAATCGACTAATTT 360

QY 441 TATGCAACAATTGCTTGATTGACAACTCTGATAGACAGAAATTCCTGGACAAAAGGCT 500
DB 361 TATGCAACAATTGCTTGATTGACAACTCTGATAGACAGAAATTCCTGGACAAAAGGCT 420

QY 501 GTGTCGTATGGATAACAGAGCTCAGTGGTTTCAGGGAAAAGTACTCTTGATGTGCACTGA 560
DB 421 GGTGTCGTATGGATAACAGAGCTCAGTGGTTTCAGGGAAAAGTACTCTTGATGTGCACTGA 480

QY 561 GTGTCGTATGGATTTCAGAGGCCACCTCACGTATGTACTTGTAGGTGACAACTTCAGAC 620
DB 481 GTGTCGTATGGATTTCAGAGGCCACCTCACGTATGTACTTGTAGGTGACAACTTCAGAC 540

QY 621 ATGGCCCTTAATAGATTTAAGCTTTAAGGCGAGAGACCGTGCAGAAATATACGAAG 680
DB 541 ATGGCCCTTAATAGATTTAAGCTTTAAGGCGAGAGACCGTGCAGAAATATACGAAG 600

QY 681 TTGGTGAATGCGCAAGCTTTTGTGTCATGCTGGGTGCATATGTCATGCTAGCTTGATAT 740
DB 601 TTGGTGAATGCGCAAGCTTTTGTGTCATGCTGGGTGCATATGTCATGCTAGCTTGATAT 660

QY 741 CTCCATACAGGAGATCGTGTATGATGCGGCTCTACTTCCACATCTTAACCTTTATTG 800

DB 661 CTCCATACAGGAGATCGTGATCGATCGCGCTCTACTTCCACATTTAACTTTATTG 720

RESULT 2
BE357876
LOCUS
DEFINITION
BE357876 675 bp mRNA linear EST 20-JUL-2000
SORGHUM bicolor cDNA, mRNA
SEQUENCE
BE357876
VERSION
BE357876.1 GI:9299433
KEYWORDS
SOURCE
ORGANISM
SORGHUM bicolor (sorghum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
1 (bases 1 to 675)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt,L.H.
An EST database from Sorghum: dark-grown seedlings
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: PolyTMix
High quality sequence start: 30
High quality sequence stop: 638
POLYA=No.

FEATURES
source
1. 675
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/note="Organ: 5-day-old dark-grown seedlings; Vector: lambda Zap; Site 1: XhoI; Site 2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

ORIGIN

Query Match 48.5%; Score 590; DB 2; Length 675;
Best Local Similarity 93.6%; Pred. No. 2.8e-137;
Matches 640; Conservative 0; Mismatches 35; Indels 9; Gaps 2;

QY 466 ATCTGATAGACAGAAATTCCTGGACAAAAGGCTGTCGTATCGATAACAGGACTCAG 525
DB 1 ATCTGATAGACAGAAATTCCTGGACAAAAGGCTGTCGTATCGATAACAGGACTCAG 60

QY 526 TGGTTCAGGAAAAAGTACTCTTTCATGTGCACTGAGTCGTGCTTCATTCGACAGGCCA 585
DB 61 TGGTTCAGGAAAAAGTACTCTTTCATGTGCACTGAGTCGTGCTTCATTCGACAGGCCA 120

QY 596 CCTCACGTCATCTTGTATGTGTGCAACCTCAGACATGGCCCTAAATAGAGATTTAAGCTT 645
DB 121 CATCACGTCATCTTGTATGTGTGCAACCTTTCAGACATGGCCCTCAATCGAGATTTAAGCTT 180

QY 646 TAAGCGAAGACCGCTGCAGAAAATATACGAAGAGTTGGTGAAGTGGCAAGACTTTTTCG 705
DB 181 TAAGCGAAGACCGCTGCAGAAAATATACGAAGAGTTGGTGAAGTGGCAAGACTTTTTCG 240

QY 706 TGATGCTGGTTCATATGTCATTTGCTAGCTTGCATATCTCCATACAGGAGATCGTGATGC 765
DB 241 CGATGCTGGTTCATATGTCATTTGCTAGCTTGCATATCTCCATACAGGAGATCGTGATGC 300

QY 766 ATGCGGTGCTCTACTTCCACATCTTAACCTTTTATGAAGTATTTATTGATTTGCCCTTAA 825

[illegible][illegible]

ORIGIN

Query Match

47.9%; Score 583.4; DB 2; Length 691;

Best Local Similarity		92.8%	Pred. No. 1.3e-135;	Matches 636; Conservative		0;	Mismatches 41;	Indels 8;		Gaps 2;
Qy	473	AGACAGAAATTCGTGGGACAAAAAGGCTGTGCTATGGATAACAGAGCTCAGTGGTTCA	532							
Db	1	AGACAGAAGTTGCTGGGACAAAAAGGCTGTGCTGTGATACAGGACTCAGTGGTTCA	60							
Qy	533	GGGAAAGTACTCTTTGATGTGCACTGAGTCGTGAGTTGCAATTCGAGAGCCACCTCAAG	592							
Db	61	GGGAAAGTACTCTTTGCTTGTGCACTGAGTCATGAGTTGCAATATAGAGGCCACATCAAG	120							
Qy	593	TATGTACTTCATCGTGTGACAACTTCAGACATGGCTAAATAGAGATTAAAGCTTTTAAGGCA	652							
Db	121	TATGTACTTCATCGTGTGACAACTTTAGACATGGCTCAATCGAGATTTAAGCTTTTAAGGCA	180							
Qy	653	GAAGACCGTGCAGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATGCT	712							
Db	181	GAAGACCGTGCAGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCGCGATGCT	240							
Qy	713	GGTGTCAATGCAATTCGTAGCTTGATATCTCCATACAGAGAGATCGTGAATGATGCCGT	772							
Db	241	GGTATCATATGCAATTCGTAGCTTGATATCTCCATACAGGAGAGATCGTGAATGATGCCGT	300							
Qy	773	GCTCTACTTCCACATTTCTAACTTTATTGAAGTATTTATTGATTTGCCCTTAAAAAATTTGT	832							
Db	301	GCTCTACTTCCAGATTTCTAACTTTATTGAAGTATTTATTGATTTGCCCTTAAAAAATTTGT	360							
Qy	833	GAAGCTCGTGATCCTAAAGGCCCTATACAAAGCTTGCACGTACAGAAAGATTAAGAGTTTC	892							
Db	361	GAAGCTCGTGATCCTAAAGGCTGTACAAGCTTGCACGCAAGAAAGATTAAGAGTTTC	420							
Qy	893	ACTGGAAATGATGATCCATACGAACCAACCAATTAATGGTGAGATAGTAAATTAAGATGAAA	952							
Db	421	ACTGGAAATGATGATCCATACGAACCGCCAGTTAATGGTGAGATAGTAAATTAAGATGAAA	480							
Qy	953	GATGAGGAATGCCCTTCAACCAAGCAATGGCCAAAGCTTCTATGCTACCTTTGAAGAA	1012							
Db	481	GATGGGAATGCCCTTCAACCAAGCAATGGCCAAAGCTTCTATCATACCTTTGAAGAG	540							
Qy	1013	AACGGATATTTGCAAGCTTAGTATATGTAATTTTGAAGAGATTTGATCTGTGTTGTGT	1072							
Db	541	AACGGATATTTGCAAGCTTAGCAAT- - -ATTTGAGAATATTGATCTGATTTCTGTGTGT	597							
Qy	1073	CCATTACTTGTGACACAATAAGATCTGTGTTGTTGGTCACATGAATAAAGGGCATCAACAT	1132							
Db	598	CCATTACTTGTGACACAATAATGATCTGTG- - -GCCATGAATAAAGGCATCAACAC	652							
Qy	1133	GTAGGAAGTAACAGAAGGTACGGTT	1157							
Db	653	ATAGCAAGTAACAGAGCGCAGTTT	677							

RESULT 4	
CK205911/c	
LOCUS	1075 bp mRNA linear EST 08-DEC-2003
DEFINITION	FGAS0107472 Triticum aestivum FGAS: Library 5 GATE 7 Triticum aestivum cDNA, mRNA sequence.
ACCESSION	CK205911
VERSION	CK205911.1 GI:39568301
KEYWORDS	EST.
SOURCE	Triticum aestivum (bread wheat)
ORGANISM	Triticum aestivum
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum. 1 (bases 1 to 1075) Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Erick,M., Gautet,D., Genswein,B., Graf,R., Gulick,P., Hryan,L.D., Laroché,A., Links,M.G., McCarthy,E.L., Monroy,A., Muszak,I., Nilson,D., Penniket,C., Roach,J.L. and Sathan.F. Functional Genomics of Abiotic Stress In Wheat and Canola Crops Unpublished (2003)
TITLE	
JOURNAL	
COMMENT	Contact : Wm L Crosby

Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: fgas_est@cs.usask.ca
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [9,754].
Plate: L5B004 row: J column: 15.

FEATURES

source

1..1075
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
/note="Vector: pCMV.SPORT6; Crown and developmental stages
of spike formation in wheat cultivar Norstar. 4 mRNA
populations were combined before constructing the library.
The first mRNA population is from 1cm crown sections after
30 days of cold acclimation. The second is from 1cm crown
sections after 11 days of deacclimation (before
deacclimation plants were fully vernalized for 49 days).
The third is from different developmental stages of spike
formation (5 to 50mm) that still have not emerged from the
leaf (dissection required). The last is from different
developmental stages of spike and seed formation after
having emerged from the leaf (visible). First strand
synthesis in this library was done in the presence of
methylated dCTP thereby protecting from internal cleavage
with NotI."

ORIGIN

Query Match 40.7%; Score 495.8; DB 7; Length 1075;
Best Local Similarity 79.9%; Pred. No. 1.5e-113;
Matches 635; Conservative 0; Mismatches 152; Indels 8; Gaps 4;
QY 334 GTGGGTGGCGCGCGCGGAATGGAGCAGCGCCGCGGGAGGCCCGCAGCAGCCAGT 393
DB 1021 GGGGTGCGCGGGAACCGGATGGAAACCCCTCAGGACGGGGCCCTTCGAAGCCCA 962
QY 394 GAAGGAGAGCCTGTAAATGCAACATGGGAAATCGACTAATATTTATGGCAATG 453
DB 961 AGTGGAAAAGCCTTAAATCCACATGGG-AATCCATTAATTTATGGCATGCTG 903
QY 454 CTTGATTGGCAATCTGATAGACAGAAATTCCTGGGACAAAAGCCTGTGCTATGGAT 513
DB 902 CCCAATTGGACAACTGAGCGACAGAAATTCCTGGGGCCAAAGGATGTTTCATTTGGAT 843
QY 514 AACAGGACTCAGTGGTTTCAGGAAAAGTACTCTTCGATGTGCATGCTGAGTTGCA 573
DB 842 CACAGGACTCAGCGTTTCAGGAAAAGTAC-CTTCCTGTGCATGGA-TCGGGGATTACA 785
QY 574 TTGCAGAGGCCACCTCACGTATGTACTTGTATGCTGATGCTGACACCTCAGACATGGCCTAAATAG 633
DB 784 CTACAGAGGCCACCAACAGTATGTTTGAATGGTGAACACCTCAGACATGGCCTCAATCG 725
QY 634 AGATTTAAGCTTTAAGGCAAGAGACCGTGCAGAAAATATACGAAGAGTTGTTGAAGTGGC 693
DB 724 AGATCTAAGCTTTCAAGGCAAGAGACCGTAGAGAAAATATACGAAGAGTTGTTGAAGTGGC 665
QY 694 AAAGCTTTTGTGATGCTGGTGTATGATGCTAGCTGATATCTCCATACAGAG 753
DB 664 AAAGCTTTTGTGAGATGCTGGTACCATATGCAATGCTAGTTGATATCTCCATACAGAG 605
QY 754 AGATCGTATGATGCTGGTGTCTACTTCCACATTTCTAACTTTATTTGAAGTATTTATGA 813
DB 604 AGACCGTATGATGCTGGGAGCTTACTTCCAGATTTCTAGATTTATTTGAAGTATTTATGA 545
QY 814 TTTGCCCTTAAAAATTTTGAAGCTGCTGATGCTTAAAGCCCTATACAAAGCTTGGCAGTAC 873

Db 544 TTTGCCACTAGAAATTATGTGAAGCTCGTGAATCTCTAAAGGGTTATACAAAGCTTGCACGCAC 485
QY 874 AGGAAGATTAAAGGTTTCACTGGAATGTATGATCCATACCAACCAATTAATGTTGA 933
DB 484 AGGAAGATTAAAGGTTTCACTGGAATGTATGATCCATACCAACCAATTAATGTTGA 425
QY 934 GATAGTAAATTAAGATGAAAGATGAGGAATGCCCTTCAACCCAAAGCAATGGCCAAAGCAAGT 993
DB 424 GATAGTAAATTAAGATGGAAGTGGGAATGCCCTTCAACCGAAGCAATGGCCCAAGCAAGT 365
QY 994 TCTATGCTACCTTCAAGAAACCGATATTTGCAAGCTTAGTATATGTAATTTTGAAG 1048
DB 364 TCTATCTTACCTTGAGAAGATGATATTTTCAGGCTTAGCACACATCTCCAGATCCAG 305
QY 1049 AAGATTGATCGATCTGTGTGTCATTTCTGTCGATCTGTCGACACATAAGATCTGTTGTTGTT 1108
DB 304 AAGATTGATCTTATTTCTGTCGATCTGTCGATCTGTCGACACATGATCAATCAATTTGTTG 245
QY 1109 CACATGAATAAAGG 1123
DB 244 CATCCGGAATAAAG 230
RESULT 5
CD863007 617 bp mRNA linear EST 11-JUL-2003
LOCUS AZ01.105F21F010130 AZ01 Triticum aestivum cDNA clone AZ01105F21,
DEFINITION mRNA sequence.
ACCESSION CD863007
VERSION CD863007.1 GI:32546823
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.
1 (bases 1 to 617)
Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiogen.fr>).
location/Qualifiers
1..617
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZ01105F21"
/tissue type="leaf"
/clone_lib="AZ01"
FEATURES
source
1..617
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZ01105F21"
/tissue type="leaf"
/clone_lib="AZ01"
ORIGIN
Query Match 40.0%; Score 486.6; DB 6; Length 617;
Best Local Similarity 88.0%; Pred. No. 2.9e-111;
Matches 543; Conservative 0; Mismatches 69; Indels 5; Gaps 1;
QY 476 CAGAAATTCCTGGGACAAAAGGCTGCTCGTATGATGATAACAGGACTCAGTGGTTTCAGGG 535
DB 1 CAGAAATTCCTGGGACAAAAGGATGTCATATGATGATCAGGACTCAGCGGTTTCAGGG 60
QY 536 AAAAGTACTCTTGCATGTGCATGCTGAGTTCGATTCAGAGGCCCACTCAGATAT 595
DB 61 AAAAGTACCGTTGCTGTGCACTGAGTCGGGAATTTACACTACAGAGGCCCACTCAGAT 120
QY 596 GTACTTGTGTGACAACTCAGACATGGCCCTAAATAGAGATTTAAGCTTTAAGGCAGNA 655

QY 603 ATGGTGCAACCTCAGACATGCGCTAAATAGAGATTAAAGCTTTAAAGCGAGAACCGTG 662
 |||||
 Db 400 ATGGTGACAACTCAGACATGCGCTAAATAGAGATTAAAGCTTTAAAGCGAGAACCGTG 459
 |||||
 QY 663 CAGAAAATATACGAAGAGTTGGTGAAGTGGCAAAAGCTTTTGGCTGATGCTGGTGTCAAT 722
 |||||
 Db 460 CAGAAAATATACGAAGAGTTGGTGAAGTGGCAAAAGCTTTTGGCTGATGCTGGTGTCAAT 519
 |||||
 QY 723 GCATTGCTAGCTGTATATCTCATACAGGAGATCGTGATGCATGCCGTGCTCTACTTC 782
 |||||
 Db 520 GCATTGCTAGCTGTATATCTCATACAGGAGATCGTGATGCATGCCGTGCTCTACTTC 579
 |||||
 QY 783 CACATCTTAACCTTATTTGAAGTATTTATTTGATTCGCCCTAAATTTGTGAAGCTCGTG 842
 |||||
 Db 580 CACATCTTAACCTTATTTGAAGTATTTATTTGATTCGCCCTAAATTTGTGAAGCTCGTG 639
 |||||
 QY 843 ATCCTTA 848
 |||||
 Db 640 ATCCTTA 645

RESULT 7
 CA204452 614 bp mRNA linear EST 25-SEP-2003
 LOCUS SCAGFL1086H06.9 FL1 Saccharum officinarum cDNA clone SCAGFL1086H06
 DEFINITION 5' mRNA sequence.

ACCESSION
 VERSION CA204452.1 GI:35241732
 KEYWORDS EST.
 SOURCE Saccharum officinarum
 ORGANISM Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.

REFERENCE
 1 (bases 1 to 614)
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
 The libraries that made SUCEST
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089

Email: parruda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bccccenter.fcav.unesp.br
 Plate: 086 row: H column: 06
 Seq primer: T7 Promoter Primer.

Location/Qualifiers
 1..614

FEATURES
 source
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCAGFL1086H06"
 /lab_host="DH10B"
 /clone_lib="FL1"

/note="Organ: Inflorescence at beginning of development (1cm-long); Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Inflorescence at beginning of development (1cm-long)]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Query Match 38.5%; Score 468; DB 6; Length 614;
 Best Local Similarity 90.5%; Pred. No. 1.4e-106;
 Matches 550; Conservative 0; Mismatches 40; Indels 18; Gaps 4;
 QY 131 CTCGCGGTCATCTCGTAAATCCACAGCGCGGCTCCCGTCTCCAGGCGCTCACCCCT 190
 |||||
 Db 15 CGCGCGGTCGGCCCGTAAATCCATAGCGCGGCTCCCGTCTCCCTCCCTCCCTCCCTCC 69
 |||||
 QY 191 AGCGATGCGGCACATCCCGGCGCTCGTGATCATGGCCTCACTCCCGTCTCTCACACTCT 250
 |||||
 Db 70 CACGATGCTCGCTGCGGCGCTCGTGATCATGGCCTCGCTCCCGTCTCTCACACGCT 129
 |||||
 QY 251 TCCGCGGCTCTCGGCAGTGATAGTGGCGCCCGCAGAGGGG-----AGGCGCGCGGT 301
 |||||
 Db 130 CCAGCGGCTCTCGGCAGCGGTGGTGGCGCCGCGAGGGGAGAGAGCGAGCGCGCGGT 189
 |||||
 QY 302 CGCGTACGCACTGCGCACCGCGGCAATTGGGCGGTGGGTGCGGCGGCGCGCGGCGGCGG 361
 |||||
 Db 190 CGCGTGGCACTGCGCACCGCGGCAATTGGGCGGTGGAT---GCGGTGGCGCGGAAATGGA 246
 |||||
 QY 362 GCAGCGCC-CGGGGAGGCCCGCACAGCCAGTGAAGGAGAGCGCTGTAATGTGGAACAT 420
 |||||
 Db 247 GCAGCGCCACGGGGAGGCCACACAGCGCCAGTGAAGGAGAGCGCTGTAATGTGGAACAT 306
 |||||
 QY 421 TGGGAAATCGCACTAATATTTTATGGCAATTTGCTTGAATGGCAATCTGTATAGACAGAA 480
 |||||
 Db 307 TGGGAAATCACTAATATTTTATGGCAATATTTGCCGATTTGGCAATCTGTATAGACAGAA 366
 |||||
 QY 481 ATTGCTGGGACAAAGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 |||||
 Db 367 GTTGCTGGGACAAAGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 426
 |||||
 QY 541 TACTCTTCATGTCACCTGAGTCTGAGTTCATTTGCGAGGCGCCACCTCAGCTATGTACT 600
 |||||
 Db 427 TACTCTTCATGTCACCTGAGTCTGAGTTCATTTGCGAGGCGCCACCTCAGCTATGTACT 486
 |||||
 QY 601 TGATGGTGACAACTCAGACATGGCCTAAATAGAGATTTAAGCTTTAAGGCGAGAGACCG 660
 |||||
 Db 487 TGATGGTGACAACTCAGACATGGCCTAAATAGAGATTTAAGCTTTAAGGCGAGAGACCG 546
 |||||
 QY 661 TGCAGAAATATACGAAGAGTTGGTGAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
 |||||
 Db 547 TGCAGAAATATACGAAGAGTTGGTGAAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 606
 |||||
 QY 721 ATGCATTG 728
 |||||
 Db 607 ATGCATTG 614

RESULT 8
 CN125017 799 bp mRNA linear EST 01-APR-2004
 LOCUS RH0H1_8_H03.bl_A002 Acid- and alkaline-treated roots Sorghum
 DEFINITION bicolor cDNA clone RH0H1_8_H03_A002 3', mRNA sequence.

ACCESSION
 VERSION CN125017.1 GI:45949036
 KEYWORDS EST.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE
 1 (bases 1 to 799)
 Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C.,
 Sun,F., Sullivan,R., Lim,S., Eastman,A. and Pratt,L.H.
 An EST database from Sorghum: acid- and alkaline-treated roots

Unpublished (2003)
 Other ESTs: RH0H1_8_H03.g1_A002
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860

Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
 Seq primer: Sug3-14 (TAGTCTAGCGCGCGGACCC)
 POLIA=yes.

FEATURES

Location/Qualifiers
 1..799
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /cultivar="BTx623"
 /db_xref="taxon:4558"
 /clone="RH01.8 H03 A002"
 /lab_host="DH10B-T1 phage-resistant E. coli"
 /clone_lib="Acid- and alkaline-treated roots"
 /note="Organ: Root; Vector: pME18S-FL3; Site.1: XhoI; Site.2: XhoI; The library was prepared from polyA+ RNA from 8-day-old roots harvested from BTx623 sorghum seedlings grown in hydroponic culture. HCl was added to a pH of 3.0 to some seedlings, KOH to a pH of 9.0 for others. Roots were harvested 3, 12 and 27 hr after addition and pooled for RNA preparation. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

ORIGIN

Query Match 37.4%; Score 455.4; DB 7; Length 799;
 Best Local Similarity 90.4%; Pred. No. 2.1e-103;
 Matches 511; Conservative 0; Mismatches 46; Indels 8; Gaps 2;

Qy 626 CTAATAGAGATTAAAGCTTTTAAGCGAGAACCGTGCAGAAAATATACGAGAGTTGGT 685
 Db 1 CTCATCGAGATTAAAGCTTTTAAGCGAGAACCGTGCAGAAAATATACGAGAGTTGGT 60

Qy 686 GAAGTGGCAAGCTTTTCTGATGCTGGTGTCTATGTCATTGCTAGCTTGATATCTCCA 745
 Db 61 GAAGTGGCAAGCTTTTTCGCGATGCTGTATCATATGTCATTGCTAGCTTGATATCTCCA 120

Qy 746 TACAGAGAGATCGTGATCGATCGCGTGTCTACTTCCACATCTCAACTTTATGAAGTA 805
 Db 121 TACAGAGAGATCGTGATCGATCGCGTGTCTACTTCCAGATTTCTAACTTTATTGAAGTA 180

Qy 806 TTTATTGATTTCCCTTAAATTTTGTGAAGCTCGTGATCCTTAAAGGCTATACAGCTT 865
 Db 181 TTTATTGATTTCCCTTAAATTTTGTGAAGCTCGTGATCCTTAAAGGCTATACAGCTT 240

Qy 866 GCACGTACAGGAAGATTAAAGGTTTCACTGGAATTGATGATCCATACGAACCCCAATT 925
 Db 241 GCACGCACAGGAAGATTAAAGTTTCACTGGAATTGATGATCCATACGAACCCCAATT 300

Qy 926 AATGTGAGATAGTAATTAAGATGAAGATGAGGAATGCCCTTACCCCAAGCAATGGCC 985
 Db 301 AATGTGAGATAGTAATTAAGATGAAGATGAGGAATGCCCTTACCCCAAGCAATGGCC 360

Qy 986 AAGCAAGTTCTATGCTTACCTTCAAGAAACGATATTTGCRAGCTTAGTATGATTTT 1045
 Db 361 AAGCAAGTTCTATGCTTACCTTCAAGAAACGATATTTGCRAGCTTAGTATGATTTT 417

Qy 1046 GAGAAGATTGATCTGATCTTGTGTGTCATTACTTTGTGGACACAAATAGATCTGTGTT 1105
 Db 418 GAGAATATTGATCTGATCTTGTGTGTCATTACTTTGTGGACACAAATAGATCTGTGTT 475

Qy 1106 GGTCAATGAATAAAGGATCAACATGTAGGAAGTAAACAGAGGTACGGTTTCATTCAGA 1165
 Db 476 ---GCCATGAATAAAGGATCAACATAGCAAGTAAACAGAGGTACGGTTTCATTCAGA 532

Qy 1166 AACGATATGGATTTCATTCGTTTAA 1190
 Db 533 ACGGAATACAGATTCATTCGTTTCA 557

RESULT 9
 CA766480
 LOCUS
 DEFINITION
 AF53-Rpf.11.K01.T7.013.abl.IRRI Drought Stress Panicle Library
 Oryza sativa (indica cultivar-group) cDNA clone C0004081 5' similar to Adenylisulfate kinase, chloroplast precursor (bc 2.7.1.25) (APS kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-phosphotransferase), mRNA sequence.
 CA766480
 CA766480.2 GI:27548315
 EST.
 SOURCE
 ORGANISM
 Oryza sativa (indica cultivar-group)
 Oryza sativa (indica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE
 1 (bases 1 to 809)
 AUTHORS
 Bennett,J., Arumugam,K., Lafitte,R., Wen,J., Rudd,S. and Bruskiewicz,R.M.
 TITLE
 IRRI Drought Stress Panicle cDNA Library
 JOURNAL
 Unpublished (2002)
 COMMENT
 On Dec 2, 2002 this sequence version replaced gi:25995735.
 Contact: Richard Bruskiewicz
 Biometrics and Bioinformatics Unit
 International Rice Research Institute
 DAPO 7777, Metro Manila, Philippines
 Tel: +63-2-845-0563
 Fax: +63-2-845-0606
 Email: r.bruskiewicz@cgiar.org
 International Rice Information System (IRIS);
 http://www.iris.irri.org: D0204080
 Assignment of putative function to the sequence by S. Rudd of the Munich Information Center for Protein Sequences
 (http://mips.gsf.de)
 Plate: 11 row: K column: 01.
 Location/Qualifiers
 1..809
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="IR64"
 /db_xref="taxon:39946"
 /clone="C0004081"
 /tissue_type="Panicles"
 /dev_stage="Flowering"
 /clone_lib="IRRI Drought Stress Panicle Library"
 /note="Vector: pBluescript II SK+; Water stress was applied by not watering for 4 consecutive days. Panicles were collected from control (well watered) and stressed plants at 2 days before heading, at heading, 50% flowering and 4 days after 50% flowering."

ORIGIN
 Query Match 36.7%; Score 446.8; DB 6; Length 809;
 Best Local Similarity 86.5%; Pred. No. 3.1e-101;
 Matches 517; Conservative 0; Mismatches 77; Indels 4; Gaps 2;

Qy 532 AGGGAAGTACTCTTGCATGTCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 591
 Db 2 AGGGAAGTACTCTTGCATGTCAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 61

Qy 592 GTATGTCATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 651
 Db 62 TTAGTCTCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 121

Qy 652 AGAAGACCGTCGAGAAAATATACGAGAGTTGGTGAAGTGGCAAGCTTTTTCGTCGTCGTC 711
 Db 122 AGAAGACCGTCGAGAAAATATACGAGAGTTGGTGAAGTGGCAAGCTTTTTCGTCGTCGTC 181

```
QY 712 TGGTGTCTATGTCATTTGCTAGCTTGATCTCTCCATACAGGAGAGATCGTGATGATCGCG 771
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 TGGATCATATGTCATTTGCTAGTTTGATATCTCCATATAGGAGAGACCGTGATCGTGCG 241
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 772 TGCTTACTTCCATCTTCACTTTATTTGAAAGTATTTATTTGATTTGCCCTTAAATAATTG 831
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 TGTGCTGCTTCCAGAGTCTAGATTTATTTGAAGTATTTATTTGATTTGCCCTAGAAATTTG 301
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 832 TGAAGCTGCTGATCTTAAGCGCTATACAGCTTGCACGTACAGAAAGATTAAGGTTT 891
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 TGAAGCTGCTGATCTTAAAGGCTATACAAAGCTTGCACGCTCTGAAAGATTTAAAGGCTT 361
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 892 CACTCGAATTTGATGATCCATACCAACCACTTAATGTTGAGATAGTAATTAAGATGAA 951
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 CACTCGAATTTGATGATCCATATGATCACCAGTGAATAGTAGTAGTAATTAAGATGGT 421
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 952 AGATGAGGAATGCCCTTCCACCAAGCAATGSCCAAGCAAGTTCTATGCTACCTTTGAAGA 1011
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 AGATGGGGAATGCTCTTACCACCAAGCAATGGCCCGCAGCATGTTCTGCTACCTTGAAGA 481
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1012 AAACGGATATTTGCAAGCTTAGTATATGATTTTGAAGATTTGATCTGTTGTTG 1071
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 GAACGGATATTTGCAAGCTTAGCATAT--ATTGTGAAGATTTGATTTCTCTGAGTG 538
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1072 TCCATTACTTTGGGACACATAAGATCTGTTGTTGTCATGATCCATTTGATTTCCA-GAATAAGATGCATCAA 1129
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 539 TCTCCACTCGTGGACATAGTGTGATCCATTTGATTTCCA-GAATAAGATGCATCAA 595
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
LOCUS CR285854
DEFINITION CR285854 Oryza sativa library (Han B) Oryza sativa cDNA clone
Y625e05p5, mRNA sequence.
ACCESSION CR285854
VERSION CR285854.1 GI:44672420
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 860)
Han,B., Feng,Q., Huang,Y.C., Ying,K., Li,Y., Guan,J.P., Zhu,J.J.,
Zhao,Q., Hu,X., Liu,Y.L., Mu,J., Yu,Z., Chen,L., Fan,D.L.,
Weng,Q.J., Zhang,L., Lu,Y.Q., Yu,S.L., Liu,X.H., Lu,T.T.,
Zhang,Y.J., Lu,Y., Li,C., Li,T., Zhang,Y., Hu,H., Jia,P.X.,
Zhang,L., Lan,L.F., Chen,W., Wu,S.A. and Xue,Y.B.
Rice cDNA EST clone
Unpublished (2003)
Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences
500# Cao Bao Road, Shanghai 200233, China
Email: bhan@ncgr.ac.cn
Clone requests: bhan@ncgr.ac.cn
This is rice cDNA est clone
Web site: http://www.ncgr.ac.cn.
location/Qualifiers
1..860
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="y625e05p5"
/clone_lib="Oryza sativa library (Han B)"

FEATURES
source
1..860
Query Match 36.3%; Score 442.2; DB 7; Length 860;
Best Local Similarity 83.2%; Pred. No. 4.6e-100;
Matches 515; Conservative 0; Mismatches 103; Indels 1; Gaps 1;

QY 345 GCGGGGGCGGAATGAGCAGCGCGCGGGAGGCCCGCAGCAGCCAGTGAAGGAGAAGC 404
```

```
Db 30 GCGGGGGAGCAGCAGCAGCAGCTCGCGGGGACGGCTCCCGCAGCCAGTCTAAAGAGAAGC 89
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 405 CTGTAATGTCCAAATTGGGAAATCGACTATATATTTATGGCACAATTTGCTTGTGTCGAC 464
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 90 CTCTAGTGTCAACCATTTGGGAAATCAACCAATATACCTGTGGCACAATTTGCCCAATTGGAC 149
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 465 AATCTGATAGACAGAAATTTGCTGGGACAAAAGGCTGTGCTGATGGATAACAGGACTCA 524
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 AATCTGAGAGCAGAACTTTGTTGGGCAAAAAGGATGTGTCTATTTGGATCACAGGACTCA 209
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 525 GTGTTTCAGGAAAAGTACTCTTGCATGTGCATGAGTCGTGAGTTGCATTTGCAGAGGCC 584
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 GCGGCTCAGGAAAAGCACTCTTGCATGTGCATGAGTCGGGAGTTGCATTTGCAGCGGTC 269
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 585 ACCTCAGCTATGTACTTGTGATGGTGACACCTCAGACATGGCCTAAATAGAGATTTAAAGCT 644
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 TCCTTACTTATGTCTTGTGATGGTGACAACTTAAGGCATGGCCTAAATCGAGATCTAAGCT 329
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 645 TTAAGGCAAGAACCGTCAGAAAATATACGAAGATTGGTGAAGTGCCAAAGCTTTTGTG 704
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 TCAAGGCAAGAACCGTCAGAAAATATACGAAGATTGGGAGAGTGCCAAAGCTATTTCG 389
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 705 CTGATGCTGGTGTCTATGTCATGCTAGCTTTGATATCTCCATACAGGAGAGATCGTGATG 764
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 CAGATGCTGGAATCATATGCAATGCTAGTTGATATCTCCATATAGGAGAGACCGTGATG 449
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 765 CATCCGCTGCTCTACTCCACATTTCTAACTTTTATGAAGTATTTATGATTTGCCCTTAA 824
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 CGTCCGCTGCTGCTTCCAGAGTCTAAATTTATGAAGTATTTATGATTTGCCACTAG 509
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 825 AAATTTTGAAGCTCGTGATCCTAAAGCCTATACAGCTTGCACGTACAGGAAGATTA 884
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 AAATTTTGAAGCTCGTGATCCTAAAGCTCTATACAGCTTGCACGTCCCGGAAGATTA 569
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 885 AAGTTTCACCTGGAATTTGATGATCCATACGAAACCAACCAATTAATGTTGAGAG-TAGTAATT 943
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 CAGCTTTCACCTGGAATTTGATGATCCATATGAATCAACAGTCACTAGTGAAGTATTAATA 629
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 944 AAGATGAAGATGAGGAAT 962
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 630 GACGTGTGAGATGGGAAGT 648
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
LOCUS BE361884
DEFINITION DGI_H05.g1_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA
sequence.
ACCESSION BE361884
VERSION BE361884.1 GI:9303441
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 746)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and
Pratt,L.H.
An EST database from Sorghum: dark-grown seedlings
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: PolyTmix
```

High quality sequence start: 30
High quality sequence stop: 719
POLYA=No.

FEATURES

Location/Qualifiers
1..746
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGI)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; Site 1: XhoI; Site 2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda Zap II.
Clones to be sequenced were prepared by mass excision."

ORIGIN

Query Match 34.7%; Score 422.2; DB 2; Length 746;
Best Local Similarity 90.3%; Pred. No. 4.8e-95;
Matches 476; Conservative 0; Mismatches 43; Indels 8; Gaps 2;
QY 664 AGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGTGATGCTGGTGTCTATATG 723
Db 1 AGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGTGATGCTGGTGTCTATATG 60
QY 724 CATTCTAGCTTGTATCTCCATACAGAGAGATCGTATGATCGCGTGTCTACTTCC 783
Db 61 CATTCTAGCTTGTATCTCCATACAGAGAGATCGTATGATCGCGTGTCTACTTCC 120
QY 784 ACATTCTACTTTATTGAAGTATTATTGATTTGCCCCCTTAAATTTGTGAAGCTCGTGA 843
Db 121 AGATTCTACTTTATTGAAGTATTATTGATTTGCCCCCTTAAATTTGTGAAGCTCGTGA 180
QY 844 TCCTAAAGCCCTATACAACTTGCAGTACAGAAAGATTAAAGGTTTCACTGGAAATTGA 903
Db 181 TCCTAAAGCTCTGTACAGCTTGCAGCACAGAAAGATTAAAGGTTTCACTGGAAATTGA 240
QY 904 TGATCCATACGAACCAACCAATTAATGGTGAAGATGAATTAAGATGAAGATGAGGAATG 963
Db 241 TGATCCATACGAACCGCCAGTTAATGGTGAAGATGAATTAAGATGAAGATGAGGAATG 300
QY 964 CCCTTACCCCAAGCAATGGCCCAAGTTCTATGCTACCTTGAAGAAACGGATATTT 1023
Db 301 CCCTTACCCCAAGCAATGGCCCAAGTTCTATGCTACCTTGAAGAAACGGATATTT 360
QY 1024 GCAAGCTTAGTATATGATTTTGAAGAGATTGATCTGATCTTGTGTGTCATTACTTGT 1083
Db 361 GCAAGCTTAGACAT--ATTGAGATATTGATCTGATCTTGTGTGTCATTACTTGT 417
QY 1084 GGACACAATAAGATCTGTTGTTGTCATGAATAAAGGCAATCAACATGTAGGAAGTAA 1143
Db 418 GGACACAATAATGATCTGTTG-----GCCATGAATAAAGGCAATCAACATAGCAAGTAA 472
QY 1144 CAGAGGTACGTTTCAATTCAGAAACGGATATGGATTCACTGTTTAA 1190
Db 473 CAGAGGCGCAGTTTGTTCAGAAACGGATATACAGATTCACTGTTTCA 519

RESULT 12

CA182186
LOCUS SCBGST3108C02.g ST3 Saccharum officinarum cDNA clone SCBGST3108C02
DEFINITION 5', mRNA sequence.
ACCESSION CA182186
VERSION CA182186.1 GI:35117704
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE 1 (bases 1 to 673)
AUTHORS Vettore, A.L., da Silva, P.R., Kemper, E.L. and Arruda, P.
TITLE The libraries that made SUCEST

JOURNAL
COMMENT

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 108 row: C column: 02
Seq primer: T7 Promoter Primer.

FEATURES

Location/Qualifiers
1..673
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCBGST3108C02"
/lab_host="DH10B"
/clone_lib="ST3"
/note="Organ: Fourth apical stalk internodes of adult
plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Fourth apical
stalk internodes of adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Query Match 34.1%; Score 415; DB 6; Length 673;
Best Local Similarity 79.6%; Pred. No. 3.1e-93;
Matches 571; Conservative 0; Mismatches 100; Indels 46; Gaps 5;
QY 73 CGGCCAGCAGCAGGGCCAAACGCAAGCAACACCTCTCTCAGCCGAGCGCCACGCT 132
Db 2 CGGCCAGCAGCAGCAACACCTCTCTCTCGGTCGACGCTGACGCGGGAGCGGCCACCG 61
QY 133 CGCCGTCATCTCTGTAATCCACAGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 191
Db 62 CGCCGTCGCGCCCGTAAATCCATAGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 121
QY 192 GCGATGCGCCACTCCCGCGCTCGTGTATCCATGCGCTCACTCCCGTCTCTCTCTCTCT 251
Db 122 ACGATGCGCGCTCTGAGCTCGTGTATCCATGCTCGCCAGCGGTGCTGGCGCC--- 178
QY 252 CCGCGGGTCTCGCCAGTGTAGTGGCGCCCGGAGGGGAGGGCGCGGTGGGTAGCG 311
Db 179 -----GCGAGGGGAGGAGAGGGGCGCGGTGGGTGGCG 217
QY 312 ACTGCCACCGCGCATTTGGCGGTTGGGTGCGCGCGCGCGCGGAAATGGAGCAGCGCCG 371
Db 218 ACTGCCACCGCGCATTTGGCGGTT-----GGCGCGGNAATGGAGCAGCGCCAC 265
QY 372 GGGAGGCGCCGCA CAGCCCGCTGTAAGGAGAGAGCTGTAATGTGGAACATTTGGGAATCGA 431
Db 266 GG-----GGAGGGCCAGTGAAGGAGAGAGCTGTAATGTGGAACATTTGGGAATCAA 317
QY 432 CTAATATTTATGGCAATTTGCTTGTATTTGACAACTCTGATAGACAGAAATGCTGGGAC 491
Db 318 CTAATATTTATGGCAATTTGCTTGTATTTGACAACTCTGATCGACAGAAATGCTGGGAC 377
QY 492 AAAAAGGCTGTCTGTATGATAACAGACTCAGTGGTTTCAAGGAAAAGTACTCTTGTGAT 551
Db 378 AAAAAGGCT-TGTCTGTGGATAACAGACTCAGTGGTTTCAAGGAAAAGTACTCTTGTGAT 436
QY 552 GTGCACTGAGTGTGAGTTGCAATTCAGAGGCCACCTCAGCTATGTACTTGTGATGATGACA 611
Db 437 GTGCACTGAGTGTGAGTTGCAATTCAGAGGCCACCTCAGCTATGTACTTGTGATGATGACA 496

Qy	540	GTACTCTTTGGATGTGCACTGAGTCGTGAGTTGTCATTGCAGAGGCCACCTCAGTATGTAC	599
Db	22	GAACCGCTGCCTGTGCACTGAGTCGGGAATTACACTACAGAGGCCACCACTATGTCC	81
Qy	600	TTGATGGTGACAACCTCAGACATGGCCTAAATAGAGATTTAAGCTTTAAGGCAGAGACC	659
Db	82	TTGATGGTGACAACCTCAGACATGGCCTCAATCGAGACTAAGCTTCAAGGCAGAGACC	141
Qy	660	GTGACAGAAAATATACGAAGAGTTGGTGAAGTGGCAACGCTTTTTCGTGATGCTGGTGTC	719
Db	142	GTCACAGAAAATATACGAAGAGTTGGTGAAGTGGCAACGCTTTTTCAGATGCTGGTACCA	201

QY	720	TATGCAATTCGTAGCCTTGAATATCTCATACAGAGAGATCGTGAATGCAATGCCGCGCTTAC
Db	202	TATGCAATTCGTAGTTTGTATCTCCATACAGAGAGATCGTGAATGCAATGCCAGCTTAC
QY	780	TTCACATCTTCACTTTATTGAAGTATTTATGATTTGCCCTTAAATAATTTGTGAAGCTC

Accession	Gene	Accession	Gene
840	Qy	899	Qy
322	Db	381	Qy
900	Qy	959	Qy

[illegible]

Qy 1020 ATTTCGAAGCTTAGTATATGTATT 1043
||||| ||||| |||
Db 502 ATTTCAGGCTTAGCACATACT 525

RESULT 14					
AW922946					
LOCUS	AW922946	721 bp	mRNA	linear	EST 19-JUL-2000
DEFINITION	DG1_47_E04.g1_A002	Dark Grown 1 (DGI)	Sorghum bicolor	cdNA, mRNA	
					sequence.

ACCESSION AW922946
 VERSION AW922946.1 GI:8088771
 KEYWORDS EST.
 SOURCE
 ORGANISM Sorghum bicolor (sorghum)
 Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 REFERENCE 1 (bases 1 to 721)
 Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and
 Pratt, L.H.
 AUTHORS Unpublished (2000)
 TITLE Contact: Cordonnier-Pratt MM
 JOURNAL Laboratory for Genomics and Bioinformatics
 COMMENT The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions
 below Phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: PolyTMix
 High quality sequence start: 23

```

FEATURES
source
Location/Qualifiers
1. .721
"/organism="Sorghum bicolor"
POLYA=No.
High quality sequence stop: 703

```

/mol_type="mRNA"
 /db_xref="taxon:4558"
 /clone_lib="Dark Grown 1 (DG1)"
 /note="Organ: 5-day-old dark-grown seedlings; Vector:
 Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
 made from poly-A RNA in the cloning vector lambda ZAP II.
 Clones to be sequenced were prepared by mass excision."

ORIGIN

```

Query Match      33.6%; Score 409.2; DB 2; Length 721;
Best Local Similarity 90.1%; Pred. No. 8.9e-92;
Matches 463; Conservative 0; Mismatches 43; Indels 8; Gaps 2;

QY 677 AGAGTTGGTGAAGTGCAGAAAGCTTTTGTGATGCTGTGTGCATATGCAATGCTAGCTTG 736
Db 1 AGAGTTGGTGAAGTGCAGAAAGCTTTTGCAGATGCTGTGTATCATATGCAATGCTAGCTTG 60

QY 737 ATATCTCCATACAGAGAGATCGTATGATGATGCGGTGCTCTACTCCCAATCTAACTTT 796
Db 61 ATATCTCCATACAGAGAGATCGTATGATGATGCGGTGCTCTACTCCCAATCTAACTTT 120

QY 797 ATTGAAGTATTTATTGATTTGCCCTTAAATAATTTGTGAAGCTCGTATCCTAAAGCCTA 856
Db 121 ATTGAAGTATTTATTGATTTGCCCTTAAATAATTTGTGAAGCTCGTATCCTAAAGCCTG 180

QY 857 TACAAGCTTGCACGTACAGAAAGATTAAGGTTTCACTGGAAATTGATGATCCATACGAA 916
Db 181 TACAAGCTTGCACGTACAGAAAGATTAAGGTTTCACTGGAAATTGATGATCCATACGAA 240

QY 917 CCACCAATTAATGGTGAGATAGTAATTAAGATGAAGATGAGGAATGCCCTTCAACCAA 976
Db 241 CGCCAGTTAATGGTGAGATAGTAATTAAGATGAAGATGAGGAATGCCCTTCAACCAA 300

QY 977 GCAATGGCCCAAGCAAGTCTATGCTACCTTGAAGAAAGCGATATTTGCAAGCTTAGTAT 1036
Db 301 GCAATGGCCCAAGCAAGTCTATGCTACCTTGAAGAAAGCGATATTTGCAAGCTTAGTAT 360

QY 1037 ATGTATTTTGAAGAATGATCTGATCTTGTGTGTCCATTACTTGTGGACACAATAAGA 1096
Db 361 AT--ATTGAGAATATTGATCTGATCTTGTGTGTCCATTACTTGTGGACACAATAAGA 417

QY 1097 TCTGTGTTGGTTCATGATTAAGGCAATCAACATGTAGGAAGTAACAGAGAGTACGGT 1156
Db 418 TCTGTGTTGGTTCATGATTAAGGCAATCAACATGTAGGAAGTAACAGAGAGTACGGT 472

QY 1157 TCATTGAGAAACGGATATGATTCATTCTGTTTAA 1190
Db 473 TCTTCAAGAAACGGATATGATTCATTCTGTTTCA 506

```

RESULT 15

```

BE361874
LOCUS BE361874
DEFINITION BE361874, 699 bp mRNA linear EST 20-JUL-2000
sequence.
ACCESSION BE361874
VERSION BE361874.1 GI:9303431
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 699)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and
Pratt,L.H.
An EST database from Sorghum: dark-grown seedlings
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860

```

Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions
 below Phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: PolyTMix
 High quality sequence start: 64
 High quality sequence stop: 697
 POLYA=No.

FEATURES

source

Location/Qualifiers
 1..699
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /db_xref="taxon:4558"
 /clone_lib="Dark Grown 1 (DG1)"
 /note="Organ: 5-day-old dark-grown seedlings; Vector:
 Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
 made from poly-A RNA in the cloning vector lambda ZAP II.
 Clones to be sequenced were prepared by mass excision."

ORIGIN

```

Query Match      32.9%; Score 400.2; DB 2; Length 699;
Best Local Similarity 89.9%; Pred. No. 1.6e-89;
Matches 454; Conservative 0; Mismatches 43; Indels 8; Gaps 2;

QY 686 GAAGTGGCAAGCTTTTGTGATGCTGTGTGCATATGCAATGCTAGCTTGATATCTCCA 745
Db 1 GAAGTGGCAAGCTTTTGTGATGCTGTGTGCATATGCAATGCTAGCTTGATATCTCCA 60

QY 746 TACAGGAGATCGTGATGATGCGGTCTCTACTTCCACATTTCACTTTATGAAGTA 805
Db 61 TACAGGAGATCGTGATGATGCGGTCTCTACTTCCAGATTTCACTTTATGAAGTA 120

QY 806 TTTATTGATTTGCCCTTAAATAATTTGTGAAGCTCGTATCCTTAAGGCTTATACAGCTT 865
Db 121 TTTATTGATTTGCCCTTAAATAATTTGTGAAGCTCGTATCCTTAAGGCTTATACAGCTT 180

QY 866 GCACGTACAGAAAGATTAAGGTTTCACTGGAATTTGATGATCCATACGAACCAATTT 925
Db 181 GCACGTACAGAAAGATTAAGGTTTCACTGGAATTTGATGATCCATACGAACCAATTT 240

QY 926 AATGTTGATAGTAATTAAGATGAAGATGAGGAATGCCCTTCAACCAAGCAATGGCC 985
Db 241 AATGTTGATAGTAATTAAGATGAAGATGAGGAATGCCCTTCAACCAAGCAATGGCC 300

QY 986 AAGCAAGTTCTATGCTACCTTGAAGAAACCGATATTTGCAAGCTTAGTATGATTTT 1045
Db 301 AAGCAAGTTCTATGCTACCTTGAAGAAACCGATATTTGCAAGCTTAGTATGATTTT 357

QY 1046 GAGAAGATTGATCTGATTTGTGTCCATTACTTGTGGACACAATAAGATCTGTTGT 1105
Db 358 GAGAATATTGATCTGATTTGTGTCCATTACTTGTGGACACAATAAGATCTGTTGT-- 415

QY 1106 GGTCAATGAATAAAGGCATCAACATGTAGGAAGTAACAGAGGTACGTTTCAATCAGA 1165
Db 416 ---GCCATGAATAAAGGCATCAACATGTAGGAAGTAACAGAGGTACGTTTCAATCAGA 472

QY 1166 AACGATATGCAATTCGTTTAA 1190
Db 473 ACGGAATACAGATTCATTCGTTTCA 497

```

Search completed: September 9, 2005, 12:40:22
 Job time : 3862 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2005, 10:07:00 ; Search time 121 Seconds
(without alignments)
1096.354 Million cell updates/sec

Title: US-10-829-432-4
Perfect score: 1826
Sequence: 1 RPFHFINQTEPLVTHQPPP.....PKMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1826	100.0	343	AAV44789	Corn Aden
2	887	48.6	224	AAV44788	Corn Aden
3	858	47.0	246	AAV44792	Wheat Ade
4	852	46.7	311	AAV44791	Soybean A
5	834	45.7	208	AAG35788	Arabidops
6	817	44.7	259	AAG14477	Arabidops
7	817	44.7	272	AAG14476	Arabidops
8	816	44.7	251	AAG38348	Arabidops
9	816	44.7	251	AAG11017	Arabidops
10	816	44.7	263	AAG38347	Arabidops
11	816	44.7	263	AAG11016	Arabidops
12	816	44.7	276	AAG38346	Arabidops
13	816	44.7	276	AAG11015	Arabidops
14	816	44.7	276	ADN74375	Thale cre
15	814	44.6	236	AAG14478	Arabidops
16	809.5	44.3	237	AAG53120	Arabidops
17	777.5	42.6	293	AAV77957	Arabidops
18	777.5	42.6	293	AAG29425	Arabidops
19	757.5	41.5	252	AAG47458	Arabidops
20	757.5	41.5	305	AAG47457	Arabidops
21	757.5	41.5	310	AAG47456	Arabidops
22	752.5	41.2	252	AAG14580	Arabidops
23	752.5	41.2	305	AAG14579	Arabidops
24	752.5	41.2	310	AAG14578	Arabidops
25	617.5	33.8	161	AAG53456	Arabidops

ALIGNMENTS

RESULT 1

AAV44789
ID AAV44789 standard; protein; 343 AA.

XX AC AAV44789;

XX DT 04-MAY-2000 (first entry)

XX DE Corn Adenylsulphate kinase-2.

XX ADenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
corn; clone p0016.ctscj40rb; transgenic plant; screen; antibody.

XX OS Zea mays.

XX PN WO200004165-A1.

XX PD 27-JAN-2000.

XX PF 13-JUL-1999; 99WO-US015809.

XX PR 14-JUL-1998; 98US-0092833P.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Falco SC, Allen SM, Anderson SL;

XX DR WPI; 2000-182430/16.

XX N-PSDB; AAZ50160.

XX PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'-
phosphosulphate kinase, useful for altering expression of sulfate
assimilation protein in plants.

XX Claim 6; Page 31-32; 42pp; English.

XX The present sequence is the corn adenylsulphate kinase (APS kinase),
also known as adenosine-5' phosphosulphate kinase. This is obtained from
clone p0016.ctscj40rb, derived from corn pooled tassels shoots, p0016 cDNA
library. This has 52% sequence identity to Catharanthus roseus APS
kinase. APS kinase is a sulphate assimilation protein, that catalyses the
conversion of adenosine-5' phosphosulphate (APS) to 3'-phospho- adenosine
-5' phosphosulphate (PAPS). The nucleotide sequence is used as probes and
primers to identify, obtain and synthesize sulphate assimilation proteins
from other plants. It is also used to produce transgenic plants, that are
useful for altering the expression levels of a sulphate assimilation
protein. The APS kinase peptides are useful for producing antibodies,

Abu23720 Protein e
Ada28651 Bacterial
Ada28185 Bacterial
Abu43064 Protein e
Abp39672 Staphyloc
Ada27651 Bacterial
Adf06882 Bacterial
Abu18587 Protein e
Abu41025 Protein e
Abu49802 Protein e
Abu43993 Protein e
Adn23471 Bacterial
Adc97735 Rhiobito
Abu31160 Protein e
Abu63457 Klebsiell
Abm67698 Phototrab
Ada28567 Bacterial
Ada21702 Bacterial
Abu48145 Protein e

26 571 31.3 200 6 ABU23720
27 555 30.4 199 8 ADS28651
28 534.5 29.3 202 8 ADS28185
29 521.5 28.6 199 6 ABU43064
30 521.5 28.6 204 5 ABP39672
31 521.5 28.6 204 8 ADS05772
32 519.5 28.5 196 8 ADS27651
33 517 28.3 226 7 ADF06882
34 514.5 28.2 197 6 ABU18587
35 514.5 28.2 204 6 ABU41025
36 510 27.9 213 6 ABU49802
37 505.5 27.7 199 6 ABU43993
38 505 27.7 652 8 ADN23471
39 504.5 27.6 627 7 ADC97735
40 503 27.5 201 6 ABU31160
41 503 27.5 216 7 ABU63457
42 501 27.4 211 6 ABM67698
43 498.5 27.3 660 8 ADS28567
44 496.5 27.2 621 8 ADS21702
45 495 27.1 201 6 ABU48145

```

CC that are used to screen and isolate cDNA clones
XX
SQ Sequence 343 AA;

Query Match 100.0%; Score 1826; DB 3; Length 343;
Best Local Similarity 100.0%; Pred. No. 2e-171;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPFHFINTQTEPLVTHTQPPSPAPGASQGGQGNLTLLSPTTTLAVILVNPORAPPVLP 60
DB 1 RPFHFINTQTEPLVTHTQPPSPAPGASQGGQGNLTLLSPTTTLAVILVNPORAPPVLP 60

QY 61 LTPSDAPLPAIVHGLTPRSSHSSAGLASDSGRREGGEGRGARTHCHRGIGRWRRRRNG 120
DB 61 LTPSDAPLPAIVHGLTPRSSHSSAGLASDSGRREGGEGRGARTHCHRGIGRWRRRRNG 120

QY 121 AAPGAPHSVPKEPVMNSIGKSTNIIWHNCLIGQSDRQKLLGQKGCVVWITGLSGSGKS 180
DB 121 AAPGAPHSVPKEPVMNSIGKSTNIIWHNCLIGQSDRQKLLGQKGCVVWITGLSGSGKS 180

QY 181 TLACALSRHLHCRGHITVYLDGDNLRHGLNRDLSPKAEDRAENIRRVGEVAKLPADAGVI 240
DB 181 TLACALSRHLHCRGHITVYLDGDNLRHGLNRDLSPKAEDRAENIRRVGEVAKLPADAGVI 240

QY 241 CIASLISPYRRDRDACCRRALLPHSNFIEVFDLPKICEARDPKGLYKLARTGKIKGFTGI 300
DB 241 CIASLISPYRRDRDACCRRALLPHSNFIEVFDLPKICEARDPKGLYKLARTGKIKGFTGI 300

QY 301 DDPEPPPINGEIVIRKMDCECPSPKAMAKQVLCYLENGYLOA 343
DB 301 DDPEPPPINGEIVIRKMDCECPSPKAMAKQVLCYLENGYLOA 343

RESULT 2
AAY44788
ID AAY44788 standard; protein; 224 AA.
XX
AC AAY44788;
XX
DT 04-MAY-2000 (first entry)
XX
DE Corn Adenylsulphate kinase-1.
XX
KW Adenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
KW 3'-phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW corn; clone cen3n.pk0088.b10; transgenic plant; screen; antibody.
XX
OS Zea mays.
XX
PN WO200004165-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015809.
XX
PR 14-JUL-1998; 98US-0092833P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Falco SC, Allen SM, Anderson SL;
XX
DR WPI; 2000-182430/16.
XX
DR N-PSDB; AA250159.
XX
PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'
PT phosphosulfate kinase, useful for altering expression of sulfate
PT assimilation protein in plants.
XX
PS Claim 6; Page 29-30; 42pp; English.
XX
CC The present sequence is the corn adenylsulphate kinase (APS kinase),
CC also known as adenosine-5' phosphosulphate kinase. It is obtained from
CC clone cen3n.pk0088.b10, derived from corn endosperm, cen3n cDNA library.
CC

CC This has 70% sequence identity to Catharanthus roseus APS kinase. APS
CC kinase is a sulphate assimilation protein, that catalyses the conversion
CC of adenosine-5' phosphosulphate (APS) to 3'-Phospho-adenosine-5'
CC phosphosulphate (PAPS). The nucleotide sequence is used as probes and
CC primers to identify, obtain and synthesise sulphate assimilation proteins
CC from other plants. It is also used to produce transgenic plants that are
CC useful for altering the expression levels of a sulphate assimilation
CC protein. The APS kinase peptides are useful for producing antibodies,
CC that are used to screen and isolate cDNA clones
XX
SQ Sequence 224 AA;

Query Match 48.6%; Score 887; DB 3; Length 224;
Best Local Similarity 77.0%; Pred. No. 7.4e-79;
Matches 171; Conservative 16; Mismatches 35; Indels 0; Gaps 0;

QY 121 AAPGEAPHSVPKEPVMNSIGKSTNIIWHNCLIGQSDRQKLLGQKGCVVWITGLSGSGKS 180
DB 2 AAAAVAGISSSSSALVTSTVGKSTNIIWHECAIGKQKQERQGLLNKQKGCVVWITGLSGSGKS 61

QY 181 TLACALSRHLHCRGHITVYLDGDNLRHGLNRDLSPKAEDRAENIRRVGEVAKLPADAGVI 240
DB 62 TLACALSRHLHCRGHITVYLDGDNLRHGLNRDLSPKAEDRAENIRRVGEVAKLPADAGLV 121

QY 241 CIASLISPYRRDRDACCRRALLPHSNFIEVFDLPKICEARDPKGLYKLARTGKIKGFTGI 300
DB 122 CIASLISPYRRDRDACCRRALLPHSNFIEVFDLPKICEARDPKGLYKLARTGKIKGFTGI 181

QY 301 DDPEPPPINGEIVIRKMDCECPSPKAMAKQVLCYLENGYLO 342
DB 182 DDPEPPPINGEIVIRKMDCECPSPKAMAKQVLCYLENGYLO 223

RESULT 3
AAY44792
ID AAY44792 standard; protein; 246 AA.
XX
AC AAY44792;
XX
DT 04-MAY-2000 (first entry)
XX
DE Wheat Adenylsulphate kinase-1.
XX
KW Adenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
KW 3'-phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW wheat; clone wrl.pk0101.e2; transgenic plant; screen; antibody.
XX
OS Triticum aestivum.
XX
PN WO200004165-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015809.
XX
PR 14-JUL-1998; 98US-0092833P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Falco SC, Allen SM, Anderson SL;
XX
DR WPI; 2000-182430/16.
XX
DR N-PSDB; AA250163.
XX
PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'
PT phosphosulfate kinase, useful for altering expression of sulfate
PT assimilation protein in plants.
XX
PS Claim 6; Page 35; 42pp; English.
XX
CC The present sequence is the wheat adenylsulphate kinase (APS kinase),
CC also known as adenosine-5' phosphosulphate kinase. This is obtained from
CC clone wrl.pk0101.e2, derived from 7 day old light grown wheat root
CC

```


PR	22-JUL-1999;	99US-01450859;
PR	22-JUL-1999;	99US-01450879;
PR	22-JUL-1999;	99US-01450899;
PR	22-JUL-1999;	99US-0145192P;
PR	23-JUL-1999;	99US-0145145P;
PR	23-JUL-1999;	99US-0145218P;
PR	23-JUL-1999;	99US-0145224P;
PR	23-JUL-1999;	99US-0145226P;
PR	26-JUL-1999;	99US-0145913P;
PR	27-JUL-1999;	99US-0145918P;
PR	27-JUL-1999;	99US-0145919P;
PR	27-JUL-1999;	99US-0145951P;
PR	28-JUL-1999;	99US-0145951P;
PR	02-AUG-1999;	99US-0146386P;
PR	02-AUG-1999;	99US-0146388P;
PR	02-AUG-1999;	99US-0146389P;
PR	03-AUG-1999;	99US-0147038P;
PR	04-AUG-1999;	99US-0147204P;
PR	04-AUG-1999;	99US-0147302P;
PR	05-AUG-1999;	99US-0147192P;
PR	05-AUG-1999;	99US-0147260P;
PR	06-AUG-1999;	99US-0147303P;
PR	06-AUG-1999;	99US-0147416P;
PR	09-AUG-1999;	99US-0147493P;
PR	09-AUG-1999;	99US-0147935P;
PR	10-AUG-1999;	99US-0148171P;
PR	11-AUG-1999;	99US-0148319P;
PR	11-AUG-1999;	99US-0148341P;
PR	12-AUG-1999;	99US-0148565P;
PR	13-AUG-1999;	99US-0148684P;
PR	16-AUG-1999;	99US-0149368P;
PR	17-AUG-1999;	99US-0149157P;
PR	18-AUG-1999;	99US-0149426P;
PR	20-AUG-1999;	99US-0149722P;
PR	20-AUG-1999;	99US-0149723P;
PR	20-AUG-1999;	99US-0149929P;
PR	23-AUG-1999;	99US-0149902P;
PR	23-AUG-1999;	99US-0149930P;
PR	25-AUG-1999;	99US-0150566P;
PR	26-AUG-1999;	99US-0150884P;
PR	27-AUG-1999;	99US-0151065P;
PR	27-AUG-1999;	99US-0151066P;
PR	27-AUG-1999;	99US-0151080P;
PR	27-AUG-1999;	99US-0151303P;
PR	31-AUG-1999;	99US-0151438P;
PR	30-SEP-1999;	99US-0154039P;
PR	20-SEP-1999;	99US-0154779P;
PR	20-SEP-1999;	99US-0159303P;
PR	01-SEP-1999;	99US-0152363P;
PR	10-SEP-1999;	99US-0153070P;
PR	13-SEP-1999;	99US-0153758P;
PR	15-SEP-1999;	99US-0154018P;
PR	16-SEP-1999;	99US-0154039P;
PR	20-SEP-1999;	99US-0154779P;
PR	22-SEP-1999;	99US-0155139P;
PR	24-SEP-1999;	99US-0155486P;
PR	24-SEP-1999;	99US-0155659P;
PR	28-SEP-1999;	99US-0156458P;
PR	29-SEP-1999;	99US-0156596P;
PR	04-OCT-1999;	99US-0157117P;
PR	05-OCT-1999;	99US-0157753P;
PR	06-OCT-1999;	99US-0157865P;
PR	07-OCT-1999;	99US-0158029P;
PR	08-OCT-1999;	99US-0158232P;
PR	12-OCT-1999;	99US-0158369P;
PR	13-OCT-1999;	99US-0159293P;
PR	13-OCT-1999;	99US-0159294P;
PR	13-OCT-1999;	99US-0159295P;
PR	14-OCT-1999;	99US-0159330P;
PR	14-OCT-1999;	99US-0159331P;
PR	14-OCT-1999;	99US-0159637P;
PR	14-OCT-1999;	99US-0159638P;
PR	18-OCT-1999;	99US-0159584P;
PR	21-OCT-1999;	99US-0160741P;
PR	21-OCT-1999;	99US-0160767P;


```
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 44.7%; Score 817; DB 3; Length 259;
Best Local Similarity 60.6%; Pred. No. 7.6e-72;
Matches 163; Conservative 32; Mismatches 52; Indels 22; Gaps 4;
QY 87 LASDGRREGGR-----GARTHCHRGIGRWVRRRRRNGAAGPGEAPHPVKEK 134
DB 1 MASPKGLSDNSRSVVVVRACVMDGQSLSHNKNKGIPELKSING-----HTGOKQG 53
QY 135 PVMSNIGKSTNHLWHNCLIGOSDRKLLGQKGCYVWITGLSGGKSTLACALSRELHCRG 194
DB 54 P-LSTVGNSTNIKWHECSVEKVRQLLDQKGCYVWITGLSGGKSTLACALNQLYQKG 112
QY 195 HLTYYLDGDNLRHGLNRDLSPKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRD 254
DB 113 KLCVILDGDNVRHGLNRDLSPKAEDRAENIRRVGEVAKLFADAGIICIASLISPYRTDRD 172
QY 255 ACRALLPHSNFIEVFDLPKICEARDPKGLYKLARTGKIKGFTGIDDPYPPINGEIV 314
DB 173 ACRNLLPEGDFEVFMDVLSVCEARDPKGLYKLARAGIKGFTGIDDPYPPLNCE--I 230
QY 315 KMDEECPSPKAMAKQVLCYLEENGYLQA 343
DB 231 SLGREGGTSPIEMAQKVVGYLDNKGYLQA 259

RESULT 7
AAG14476
ID AAG14476 standard; protein; 272 AA.
XX AAG14476;
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SRQ ID NO: 14354.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0121806P.
XX 09-MAR-1999; 99US-0123548P.
```

PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151388P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.

PR	07-OCT-1999;	99US-0158029P.	PN	EP1033405-A2.
PR	08-OCT-1999;	99US-0158232P.	XX	
PR	12-OCT-1999;	99US-0158369P.	PD	06-SEP-2000.
PR	13-OCT-1999;	99US-0159293P.	XX	
PR	13-OCT-1999;	99US-0159294P.	PF	25-FEB-2000; 2000EP-00301439.
PR	14-OCT-1999;	99US-0159295P.	XX	
PR	14-OCT-1999;	99US-0159330P.	PR	25-FEB-1999;
PR	14-OCT-1999;	99US-0159331P.	PR	05-MAR-1999;
PR	14-OCT-1999;	99US-0159637P.	PR	09-MAR-1999;
PR	14-OCT-1999;	99US-0159638P.	PR	23-MAR-1999;
PR	18-OCT-1999;	99US-0159584P.	PR	25-MAR-1999;
PR	21-OCT-1999;	99US-0160741P.	PR	99US-0126264P.
PR	21-OCT-1999;	99US-0160767P.	PR	99US-0126785P.
PR	21-OCT-1999;	99US-0160768P.	PR	99US-0127462P.
PR	21-OCT-1999;	99US-0160770P.	PR	99US-0128234P.
PR	21-OCT-1999;	99US-0160814P.	PR	99US-0128714P.
PR	21-OCT-1999;	99US-0160815P.	PR	99US-0129845P.
PR	22-OCT-1999;	99US-0160980P.	PR	99US-0130077P.
PR	22-OCT-1999;	99US-0160981P.	PR	99US-0130449P.
PR	22-OCT-1999;	99US-0160989P.	PR	99US-0130510P.
PR	25-OCT-1999;	99US-0161404P.	PR	99US-0130891P.
PR	25-OCT-1999;	99US-0161405P.	PR	99US-0131449P.
PR	25-OCT-1999;	99US-0161406P.	PR	99US-0132048P.
PR	26-OCT-1999;	99US-0161359P.	PR	99US-0132407P.
PR	26-OCT-1999;	99US-0161360P.	PR	99US-0132484P.
PR	26-OCT-1999;	99US-0161361P.	PR	99US-0132485P.
PR	28-OCT-1999;	99US-0161920P.	PR	99US-0132486P.
PR	28-OCT-1999;	99US-0161992P.	PR	99US-0132487P.
PR	28-OCT-1999;	99US-0161993P.	PR	99US-0132863P.
PR	29-OCT-1999;	99US-0162142P.	PR	99US-0134256P.
Query Match 44.7%; Score 817; DB 3; Length 272;			PR	99US-0134256P.
Best Local Similarity 60.6%; Pred. No. 8.1e-72;			PR	99US-0134218P.
Matches 163; Conservative 32; Mismatches 52; Indels 22; Gaps 4;			PR	99US-0134219P.
QY	87 LASDGRREGEGR-----GARTHCHRGIGRWVRRRRRNGAAGPEAPHSVPEKEK 134	134	PR	99US-0134221P.
Db	14 MASPKGLSDNSRSVVVVRACVSMDSQTLSHKNKGSIPELKSI NG-----HTGQKQG 66	66	PR	99US-0134370P.
QY	135 PWSNIGKSTNLIWNCLIGQSDRQKLQKQKCVVWITGLSGSKSTLACALSRELHCRG 194	194	PR	99US-0134370P.
Db	67 P-LSTVGNSTNFKWHECSVEKVDQRLLDQKQCVI WVTGLSGSKSTLACALNQMLYQKG 125	125	PR	99US-0134768P.
QY	195 HLTYYLDGNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGVICTASLISPYRRDRD 254	254	PR	99US-0134941P.
Db	126 KLCYILDGNVVRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGIICIASLISPYRTDRD 185	185	PR	99US-0135124P.
QY	255 ACRALLPHSNFIEVFDLPLKICEARDPKGLYKLARTGKIKGFTGIDDPYEPPIGEIVI 314	314	PR	99US-0135353P.
Db	186 ACRNLLPEGDFVEVFMDFVLSVCEARDPKGLYKLARAGIKGFTGIDDPYEPPLNCE--I 243	243	PR	99US-0135629P.
QY	315 KMKDECPSPKMAKQVLCYLSYNGYLOA 343	343	PR	99US-0136021P.
Db	244 SLGREGGTSPIEMAEKVVGYLDNKGYLQA 272	272	PR	99US-0136392P.
RESULT 8			PR	99US-0136782P.
AAG38348			PR	99US-0137222P.
ID	AAG38348 standard; protein; 251 AA.		PR	99US-0137502P.
XX			PR	99US-0137724P.
AC			PR	99US-0138094P.
XX			PR	99US-0138540P.
DT	18-OCT-2000 (first entry)		PR	99US-0138847P.
XX			PR	99US-0139119P.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 47297.		PR	99US-0139452P.
XX			PR	99US-0139453P.
KW	Protein identification; signal transduction pathway; metabolic pathway;		PR	99US-0139453P.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		PR	99US-0139456P.
XX	termination sequence.		PR	99US-0139457P.
OS	Arabidopsis thaliana.		PR	99US-0139458P.
XX			PR	99US-0139459P.
XX			PR	99US-0139460P.
XX			PR	99US-0139461P.
XX			PR	99US-0139462P.
XX			PR	99US-0139463P.
XX			PR	99US-0139750P.
XX			PR	99US-0139763P.
XX			PR	99US-0139817P.
XX			PR	99US-0139899P.
XX			PR	99US-0140353P.
XX			PR	99US-0140354P.
XX			PR	99US-0140695P.
XX			PR	99US-0140823P.
XX			PR	99US-0140991P.
XX			PR	99US-0141287P.
XX			PR	99US-0141842P.

PR 01-JUL-1999;	99US-0142154P.
PR 02-JUL-1999;	99US-0142055P.
PR 06-JUL-1999;	99US-0142390P.
PR 08-JUL-1999;	99US-0142803P.
PR 09-JUL-1999;	99US-0142820P.
PR 12-JUL-1999;	99US-0142877P.
PR 13-JUL-1999;	99US-0143342P.
PR 14-JUL-1999;	99US-0143624P.
PR 15-JUL-1999;	99US-0144005P.
PR 16-JUL-1999;	99US-0144085P.
PR 16-JUL-1999;	99US-0144086P.
PR 19-JUL-1999;	99US-0144325P.
PR 19-JUL-1999;	99US-0144331P.
PR 19-JUL-1999;	99US-0144332P.
PR 19-JUL-1999;	99US-0144333P.
PR 19-JUL-1999;	99US-0144334P.
PR 20-JUL-1999;	99US-0144335P.
PR 20-JUL-1999;	99US-0144352P.
PR 20-JUL-1999;	99US-0144632P.
PR 21-JUL-1999;	99US-0144884P.
PR 21-JUL-1999;	99US-0144814P.
PR 21-JUL-1999;	99US-0145086P.
PR 21-JUL-1999;	99US-0145088P.
PR 22-JUL-1999;	99US-0145085P.
PR 22-JUL-1999;	99US-0145087P.
PR 22-JUL-1999;	99US-0145089P.
PR 22-JUL-1999;	99US-0145192P.
PR 23-JUL-1999;	99US-0145145P.
PR 23-JUL-1999;	99US-0145218P.
PR 23-JUL-1999;	99US-0145224P.
PR 26-JUL-1999;	99US-0145276P.
PR 27-JUL-1999;	99US-0145913P.
PR 27-JUL-1999;	99US-0145918P.
PR 27-JUL-1999;	99US-0145919P.
PR 28-JUL-1999;	99US-0145951P.
PR 02-AUG-1999;	99US-0146386P.
PR 02-AUG-1999;	99US-0146388P.
PR 02-AUG-1999;	99US-0146389P.
PR 03-AUG-1999;	99US-0147038P.
PR 04-AUG-1999;	99US-0147204P.
PR 04-AUG-1999;	99US-0147302P.
PR 05-AUG-1999;	99US-0147192P.
PR 05-AUG-1999;	99US-0147260P.
PR 06-AUG-1999;	99US-0147303P.
PR 06-AUG-1999;	99US-0147416P.
PR 09-AUG-1999;	99US-0147493P.
PR 09-AUG-1999;	99US-0147935P.
PR 10-AUG-1999;	99US-0148171P.
PR 11-AUG-1999;	99US-0148319P.
PR 12-AUG-1999;	99US-0148341P.
PR 13-AUG-1999;	99US-0148565P.
PR 13-AUG-1999;	99US-0148684P.
PR 16-AUG-1999;	99US-0149368P.
PR 17-AUG-1999;	99US-0149175P.
PR 18-AUG-1999;	99US-0149426P.
PR 20-AUG-1999;	99US-0149722P.
PR 20-AUG-1999;	99US-0149723P.
PR 20-AUG-1999;	99US-0149929P.
PR 23-AUG-1999;	99US-0149902P.
PR 23-AUG-1999;	99US-0149930P.
PR 25-AUG-1999;	99US-0150566P.
PR 26-AUG-1999;	99US-0150884P.
PR 27-AUG-1999;	99US-0151065P.
PR 27-AUG-1999;	99US-0151068P.
PR 27-AUG-1999;	99US-0151080P.
PR 30-AUG-1999;	99US-0151303P.
PR 31-AUG-1999;	99US-0151438P.
PR 01-SEP-1999;	99US-0151930P.
PR 07-SEP-1999;	99US-0152363P.
PR 10-SEP-1999;	99US-0153070P.
PR 13-SEP-1999;	99US-0153758P.
PR 15-SEP-1999;	99US-0154018P.
PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-01542154P.
PR 02-JUL-1999;	99US-0154255P.
PR 06-JUL-1999;	99US-0154290P.
PR 08-JUL-1999;	99US-01542803P.
PR 09-JUL-1999;	99US-01542820P.
PR 12-JUL-1999;	99US-01542877P.
PR 13-JUL-1999;	99US-01543342P.
PR 14-JUL-1999;	99US-01543624P.
PR 15-JUL-1999;	99US-01544005P.
PR 16-JUL-1999;	99US-01544085P.
PR 16-JUL-1999;	99US-01544086P.
PR 19-JUL-1999;	99US-01544325P.
PR 19-JUL-1999;	99US-01544331P.
PR 19-JUL-1999;	99US-01544332P.
PR 19-JUL-1999;	99US-01544333P.
PR 19-JUL-1999;	99US-01544334P.
PR 20-JUL-1999;	99US-01544335P.
PR 20-JUL-1999;	99US-01544352P.
PR 20-JUL-1999;	99US-01544632P.
PR 21-JUL-1999;	99US-01544884P.
PR 21-JUL-1999;	99US-01544814P.
PR 21-JUL-1999;	99US-01545086P.
PR 21-JUL-1999;	99US-01545088P.
PR 22-JUL-1999;	99US-01545085P.
PR 22-JUL-1999;	99US-01545087P.
PR 22-JUL-1999;	99US-01545089P.
PR 22-JUL-1999;	99US-01545192P.
PR 23-JUL-1999;	99US-01545145P.
PR 23-JUL-1999;	99US-01545218P.
PR 23-JUL-1999;	99US-01545224P.
PR 26-JUL-1999;	99US-01545276P.
PR 27-JUL-1999;	99US-01545913P.
PR 27-JUL-1999;	99US-01545918P.
PR 27-JUL-1999;	99US-01545919P.
PR 28-JUL-1999;	99US-01545951P.
PR 02-AUG-1999;	99US-01546386P.
PR 02-AUG-1999;	99US-01546388P.
PR 02-AUG-1999;	99US-01546389P.
PR 03-AUG-1999;	99US-01547038P.
PR 04-AUG-1999;	99US-01547204P.
PR 04-AUG-1999;	99US-01547302P.
PR 05-AUG-1999;	99US-01547192P.
PR 05-AUG-1999;	99US-01547260P.
PR 06-AUG-1999;	99US-01547303P.
PR 06-AUG-1999;	99US-01547416P.
PR 09-AUG-1999;	99US-01547493P.
PR 09-AUG-1999;	99US-01547935P.
PR 10-AUG-1999;	99US-01548171P.
PR 11-AUG-1999;	99US-01548319P.
PR 12-AUG-1999;	99US-01548341P.
PR 13-AUG-1999;	99US-01548565P.
PR 13-AUG-1999;	99US-01548684P.
PR 16-AUG-1999;	99US-01549368P.
PR 17-AUG-1999;	99US-01549175P.
PR 18-AUG-1999;	99US-01549426P.
PR 20-AUG-1999;	99US-01549722P.
PR 20-AUG-1999;	99US-01549723P.
PR 20-AUG-1999;	99US-01549929P.
PR 23-AUG-1999;	99US-01549902P.
PR 23-AUG-1999;	99US-01549930P.
PR 25-AUG-1999;	99US-01550566P.
PR 26-AUG-1999;	99US-01550884P.
PR 27-AUG-1999;	99US-01551065P.
PR 27-AUG-1999;	99US-01551068P.
PR 27-AUG-1999;	99US-01551080P.
PR 30-AUG-1999;	99US-01551303P.
PR 31-AUG-1999;	99US-01551438P.
PR 01-SEP-1999;	99US-01551930P.
PR 07-SEP-1999;	99US-01552363P.
PR 10-SEP-1999;	99US-01553070P.
PR 13-SEP-1999;	99US-01553758P.
PR 15-SEP-1999;	99US-01554018P.
PR 16-SEP-1999;	99US-01554039P.
PR 20-SEP-1999;	99US-01554779P.
PR 22-SEP-1999;	99US-01555139P.
PR 23-SEP-1999;	99US-0155486P.
PR 24-SEP-1999;	99US-0155659P.
PR 28-SEP-1999;	99US-0155658P.
PR 29-SEP-1999;	99US-0155659P.
PR 04-OCT-1999;	99US-0157117P.
PR 05-OCT-1999;	99US-015753P.
PR 06-OCT-1999;	99US-0157865P.
PR 07-OCT-1999;	99US-0158023P.
PR 08-OCT-1999;	99US-0158232P.
PR 12-OCT-1999;	99US-0158369P.
PR 13-OCT-1999;	99US-0159293P.
PR 13-OCT-1999;	99US-0159294P.
PR 13-OCT-1999;	99US-0159295P.
PR 14-OCT-1999;	99US-0159323P.
PR 14-OCT-1999;	99US-0159330P.
PR 14-OCT-1999;	99US-0159331P.
PR 14-OCT-1999;	99US-0159637P.
PR 14-OCT-1999;	99US-0159638P.
PR 18-OCT-1999;	99US-0159584P.
PR 21-OCT-1999;	99US-0160741P.
PR 21-OCT-1999;	99US-0160767P.
PR 21-OCT-1999;	99US-0160768P.
PR 21-OCT-1999;	99US-0160770P.
PR 21-OCT-1999;	99US-0160814P.
PR 21-OCT-1999;	99US-0160815P.
PR 22-OCT-1999;	99US-0160980P.
PR 22-OCT-1999;	99US-0160981P.
PR 22-OCT-1999;	99US-0160989P.
PR 25-OCT-1999;	99US-0161404P.
PR 25-OCT-1999;	99US-0161405P.
PR 25-OCT-1999;	99US-0161406P.
PR 26-OCT-1999;	99US-0161359P.
PR 26-OCT-1999;	99US-0161360P.
PR 26-OCT-1999;	99US-0161361P.
PR 28-OCT-1999;	99US-0161920P.
PR 28-OCT-1999;	99US-0161992P.
PR 28-OCT-1999;	99US-0161993P.
PR 29-OCT-1999;	99US-0162142P.
Query Match 44.7%; Score 816; DB 3; Length 251;	
Best Local Similarity 64.4%; Pred. No. 9.1e-72;	
Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;	
Qy 100	GARTCHRCIGRWRRRRRNGAAGEAP---HSPVKEPVMNSIGKSTNIIWHNCLIGQS 156
Db 18	GSQTLSH-----NKNGSIPEVKSINGHTGQKQGP-LSTVGNSTNIKWHECSVEKV 66
Qy 157	DRQKLGQKGCWVWITGLSGSGKSTLACALSRHLHCRGHLYVLDGDNLRHGLNRDLSPK 216
Db 67	DRQRLDQKGCVIWITGLSGSGKSTLACALNQMLYQKGLCYILOGDNVRHGLNRDLSPK 126
Qy 217	AEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACRALLPHSNFIEVFDLPK 276
Db 127	AEDRAENIRRVGEVAKLFADAGIICIASLISPYRTDRDACSLLPEGDFVEFMDVPLSV 186
Qy 277	CEARDPKGLYKLARTGKIKGFTGIDDDPYEPPIINGEIVIKMKDECPSPKAKQVLCYLE 336
Db 187	CEARDPKGLYKLARAGIKGFTGIDDDPYEPPLNCE--ISLREGGTSPTEMAEKVVGYLD 244
Qy 337	ENGYLOA 343
Db 245	NGYLQA 251
RESULT 9	
AAG11017	
ID	AAG11017 standard; protein; 251 AA.
XX	
XX	AAG11017;
XX	
DT	17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 9573.
DE Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 20-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.


```
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 44.7%; Score 816; DB 3; Length 251;
Best Local Similarity 64.4%; Pred. No. 9.1e-72;
Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;

QY 100 GARTCHRGIGRWRRRRNGAAGEAP---HSPVKEKPVMSNICKSTNHLWNCLIGOS 156
Db 18 GSGTSLSH-----NKNGSIPEVKSINGHTGQKQP-ISTVGNSTNKKHCECSVEKV 66

QY 157 DRQKLQKQGVWITGLSGSKSTLACALSRELHCRGHLTYVLGDNDLRLHGLNRDLSFK 216
Db 67 DRQLLDQKGVWITGLSGSKSTLACALNQLYQKGLCYILDDNVNRHGLNRDLSFK 126

QY 217 AEDRAENRRVGEAKLFADAGVICTASLISPYRDRDRAALLPHSNFIEVFDLPKI 276
Db 127 AEDRAENRRVGEAKLFADAGVICTASLISPYRDRDRAALLPHSNFIEVFDLPKI 186

QY 277 CEARDPKGLYKLARTKGIKGTGIDDPYEPPIINGEIVIMKDEECFSPKAKAKOVLVLYE 336
Db 187 CEARDPKGLYKLARTKGIKGTGIDDPYEPPIINGEIVIMKDEECFSPKAKAKOVLVLYE 336

QY 337 ENGYLQA 343
Db 245 NKGYLQA 251
```

RESULT 10

AAG38347

ID AAG38347 standard; protein; 263 AA.

XX

AC AAG38347;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 47296.

XX

KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR

PR 05-MAR-1999; 99US-0123180P.

PR

PR 09-MAR-1999; 99US-0123548P.

PR

PR 23-MAR-1999; 99US-0125788P.

PR

PR 25-MAR-1999; 99US-0126264P.

PR

PR 29-MAR-1999; 99US-0126785P.

PR

PR 01-APR-1999; 99US-0127462P.

PR

PR 06-APR-1999; 99US-0128234P.

PR

PR 08-APR-1999; 99US-0128714P.

PR

PR 16-APR-1999; 99US-0129845P.

PR

PR 19-APR-1999; 99US-0130077P.

PR

PR 21-APR-1999; 99US-0130510P.

PR

PR 23-APR-1999; 99US-0130891P.

PR

PR 30-APR-1999; 99US-0132048P.

PR

PR 30-APR-1999; 99US-0132407P.

PR

PR 04-MAY-1999; 99US-0132484P.

PR

PR 05-MAY-1999; 99US-0132485P.

PR

PR 06-MAY-1999; 99US-0132486P.

PR

PR 06-MAY-1999; 99US-0132487P.

PR

PR 07-MAY-1999; 99US-0132863P.

PR

PR 11-MAY-1999; 99US-0134256P.

PR

PR 14-MAY-1999; 99US-0134218P.

PR

PR 14-MAY-1999; 99US-0134219P.

PR

PR 14-MAY-1999; 99US-0134221P.

PR

PR 14-MAY-1999; 99US-0134370P.

PR

PR 18-MAY-1999; 99US-0134768P.

PR

PR 19-MAY-1999; 99US-0134941P.

PR

PR 20-MAY-1999; 99US-0135124P.

PR

PR 21-MAY-1999; 99US-0135353P.

PR

PR 24-MAY-1999; 99US-0135629P.

PR

PR 25-MAY-1999; 99US-0136021P.

PR

PR 27-MAY-1999; 99US-0136392P.

PR

PR 28-MAY-1999; 99US-0136782P.

PR

PR 01-JUN-1999; 99US-0137222P.

PR

PR 03-JUN-1999; 99US-0137528P.

PR

PR 04-JUN-1999; 99US-0137502P.

PR

PR 07-JUN-1999; 99US-0137724P.

PR

PR 08-JUN-1999; 99US-0138094P.

PR

PR 10-JUN-1999; 99US-0138540P.

PR

PR 10-JUN-1999; 99US-0138847P.

PR

PR 14-JUN-1999; 99US-0139119P.

PR

PR 16-JUN-1999; 99US-0139452P.

PR

PR 16-JUN-1999; 99US-0139453P.

PR

PR 17-JUN-1999; 99US-0139492P.

PR

PR 18-JUN-1999; 99US-0139454P.

PR

PR 18-JUN-1999; 99US-0139455P.

PR

PR 18-JUN-1999; 99US-0139456P.

PR

PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0142878P.
PR	01-JUL-1999;	99US-0142842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	06-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159283P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	18-OCT-1999;	99US-0159638P.
PR	21-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	26-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 44.7%; Score 816; DB 3; Length 263;		
Best Local Similarity 64.4%; Pred. No. 9.7e-72;		
Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;		
Qy	100	GARTCHRGIGRWVRRRRNGAAPCEAP---HSPVKPKPVMSNIGKSTNIIWHLNCLIQGS 156
Db	30	GSQTLSH-----NKGSIPEVKSINGHTGOKQGP-ISTVGNSTNIKWHECSVEKV 78
Qy	157	DRQKLLGQKGVWITGLSGSKSTLACALRELHCRGHLTYVLDGDNLRHGLNRLDSFK 216
Db	79	DRQRLLDQKGVWITGLSGSKSTLACALNQMPLYKQKLCYILDGDNVRHGLNRLDSFK 138
Qy	217	AEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACRALLPHSNFIEVFDLPKI 276

Db 139 AEDRAENIRRVGEVAKLFADAGLICIASLISPYRTDRDACRSLLPEGDFVEFMDVPLSV 198
Qy 277 CEARDPGLYKLARTGKIKGFTGIDDPYEPPIGEIVIKWKDBECPSPKAMAKQVLCYLE 336
Db 199 CEARDPGLYKLARAGKIKGFTGIDDPYEPPLNCE--ISLGREGGTSPIEMAEKVGVYLD 256
Qy 337 ENGYYLOA 343
Db 257 NKGYLQA 263
RESULT 11
AAG11016
ID AAG11016 standard; protein; 263 AA.
XX AC AAG11016;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 9572.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144335P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

PR	09-AUG-1999;	99US-0147935P.	Db	30	GSQTLSH-----NKGSIPEVKSINGHTCQKQCP-LSTVGNSTNIKWHGCSVEKV	78
PR	10-AUG-1999;	99US-0148171P.	Qy	157	DRQKLGQKGVVWITGLSGSGKSTLACALSREHLCHRGHLTYVLDDGNLRHGLNRDLSFK	216
PR	11-AUG-1999;	99US-0148319P.	Db	79	DRQRLDQKGVVWITGLSGSGKSTLACALNQMLYQKGKJCYLLDGNVRHGLNRDLSFK	138
PR	12-AUG-1999;	99US-0148341P.	Qy	217	AEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACRALLPHSNFIEVFIDLPLKI	276
PR	13-AUG-1999;	99US-0148565P.	Db	139	AEDRAENIRRVGEVAKLFADAGIICIASLISPYRTDRDACRSLLPEDGFVEFVMDVPLSV	198
PR	16-AUG-1999;	99US-0148684P.	Qy	277	CEARDPKGLYKLTARTKIKGFTGIDDPYEPPIGEIIVIMKDECPSPKAKQVLCYLE	336
PR	17-AUG-1999;	99US-0149175P.	Db	199	CEARDPKGLYKLTARAGIKGFTGIDDPYEPPLNCE--ISLREGGTSPIENAEKVVGILD	256
PR	18-AUG-1999;	99US-014926P.	Qy	337	ENGYLOA	343
PR	20-AUG-1999;	99US-0149722P.	Db	257	NKGYLQA	263
PR	20-AUG-1999;	99US-0149723P.				
PR	20-AUG-1999;	99US-0149929P.				
PR	23-AUG-1999;	99US-0149902P.				
PR	23-AUG-1999;	99US-0149930P.				
PR	25-AUG-1999;	99US-0150566P.				
PR	26-AUG-1999;	99US-0150884P.				
PR	27-AUG-1999;	99US-0150884P.				
PR	27-AUG-1999;	99US-015065P.				
PR	27-AUG-1999;	99US-0151066P.				
PR	27-AUG-1999;	99US-0151080P.				
PR	30-AUG-1999;	99US-0151303P.				
PR	31-AUG-1999;	99US-0151438P.				
PR	01-SEP-1999;	99US-0151930P.				
PR	07-SEP-1999;	99US-0152363P.				
PR	10-SEP-1999;	99US-0153070P.				
PR	13-SEP-1999;	99US-0153758P.				
PR	15-SEP-1999;	99US-0154018P.				
PR	16-SEP-1999;	99US-0154039P.				
PR	20-SEP-1999;	99US-0154779P.				
PR	22-SEP-1999;	99US-0155139P.				
PR	23-SEP-1999;	99US-0155486P.				
PR	24-SEP-1999;	99US-0155659P.				
PR	28-SEP-1999;	99US-0156458P.				
PR	29-SEP-1999;	99US-0156596P.				
PR	04-OCT-1999;	99US-0157117P.				
PR	05-OCT-1999;	99US-0157753P.				
PR	06-OCT-1999;	99US-0157865P.				
PR	07-OCT-1999;	99US-0158029P.				
PR	08-OCT-1999;	99US-0158232P.				
PR	12-OCT-1999;	99US-0158369P.				
PR	13-OCT-1999;	99US-0159293P.				
PR	13-OCT-1999;	99US-0159294P.				
PR	13-OCT-1999;	99US-0159295P.				
PR	14-OCT-1999;	99US-0159329P.				
PR	14-OCT-1999;	99US-0159330P.				
PR	14-OCT-1999;	99US-0159331P.				
PR	14-OCT-1999;	99US-0159637P.				
PR	14-OCT-1999;	99US-0159638P.				
PR	18-OCT-1999;	99US-0159584P.				
PR	21-OCT-1999;	99US-0160741P.				
PR	21-OCT-1999;	99US-0160767P.				
PR	21-OCT-1999;	99US-0160768P.				
PR	21-OCT-1999;	99US-0160770P.				
PR	21-OCT-1999;	99US-0160814P.				
PR	21-OCT-1999;	99US-0160815P.				
PR	22-OCT-1999;	99US-0160980P.				
PR	22-OCT-1999;	99US-0160981P.				
PR	22-OCT-1999;	99US-0160989P.				
PR	25-OCT-1999;	99US-0161404P.				
PR	25-OCT-1999;	99US-0161405P.				
PR	25-OCT-1999;	99US-0161406P.				
PR	25-OCT-1999;	99US-0161359P.				
PR	26-OCT-1999;	99US-0161360P.				
PR	26-OCT-1999;	99US-0161361P.				
PR	26-OCT-1999;	99US-0161920P.				
PR	28-OCT-1999;	99US-0161992P.				
PR	28-OCT-1999;	99US-0161993P.				
PR	29-OCT-1999;	99US-0162142P.				
PR	29-OCT-1999;	99US-0162142P.				
Query Match 44.7%; Score 816; DB 3; Length 263;						
Best Local Similarity 64.4%; Pred. No. 9.7e-72;						
Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;						
QY	100	GARTHCHRGIGRWRRRRRNGAAGEAP---HSPVKEKPMNSNIGKSTNLMHNCILIGOS	156			

Db	30	GSQTLSH-----NKGSIPEVKSINGHTCQKQCP-LSTVGNSTNIKWHGCSVEKV	78
Qy	157	DRQKLGQKGVVWITGLSGSGKSTLACALSREHLCHRGHLTYVLDDGNLRHGLNRDLSFK	216
Db	79	DRQRLDQKGVVWITGLSGSGKSTLACALNQMLYQKGKJCYLLDGNVRHGLNRDLSFK	138
Qy	217	AEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACRALLPHSNFIEVFIDLPLKI	276
Db	139	AEDRAENIRRVGEVAKLFADAGIICIASLISPYRTDRDACRSLLPEDGFVEFVMDVPLSV	198
Qy	277	CEARDPKGLYKLTARTKIKGFTGIDDPYEPPIGEIIVIMKDECPSPKAKQVLCYLE	336
Db	199	CEARDPKGLYKLTARAGIKGFTGIDDPYEPPLNCE--ISLREGGTSPIENAEKVVGILD	256
Qy	337	ENGYLOA	343
Db	257	NKGYLQA	263
RESULT 12			
AAG38346	AAG38346 standard; protein; 276 AA.		
ID	AAG38346		
AC	AAG38346;		
XX	18-OCT-2000 (first entry)		
DT	Arabidopsis thaliana protein fragment SEQ ID NO: 47295.		
DE	Protein identification; signal transduction pathway; metabolic pathway;		
XX	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
XX	Arabidopsis thaliana.		
OS	EP1033405-A2.		
PN	06-SEP-2000.		
PD	25-FEB-2000; 2000EP-00301439.		
PF	25-FEB-1999; 99US-0121825P.		
XX	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	30-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	04-MAY-1999; 99US-0132407P.		
PR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	06-MAY-1999; 99US-0132487P.		
PR	07-MAY-1999; 99US-0132863P.		
PR	11-MAY-1999; 99US-0134256P.		
PR	14-MAY-1999; 99US-0134218P.		
PR	14-MAY-1999; 99US-0134219P.		
PR	14-MAY-1999; 99US-0134221P.		
PR	14-MAY-1999; 99US-0134370P.		
PR	18-MAY-1999; 99US-0134768P.		
PR	19-MAY-1999; 99US-0134941P.		
PR	20-MAY-1999; 99US-0135124P.		
PR	21-MAY-1999; 99US-0135353P.		

PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159283P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160815P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.

```
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 44.7%; Score 816; DB 3; Length 276;
Best Local Similarity 64.4%; Pred. No. 1e-71;
Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;

QY 100 GARTCHRGIGRVRRRRNGAAGEAP---HSPVKEKPEVMSNICKSTNIIWHNCLIGOS 156
Db 43 GSQTLSH-----NKGSIPEVKSINGHTGOKGP-LSTGVNSTIKWHECSEVKV 91
QY 157 DRQKLLGQKGVWITGLSGSKSTLACALSRELHCRGHLTYVLDCDNLRLHGLNDRDLSFK 216
Db 92 DRQRLDDQKGVWITGLSGSKSTLACALNQLKQKGLCYLLDGNVNRHGLNDRDLSFK 151
QY 217 AEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACRALLPHSNFIEVFDLPKI 276
Db 152 AEDRAENIRRVGEVAKLFADAGIICIASLISPYRTDRDACRSLLPEDGFVEVFMVPLSV 211
QY 277 CEARDPKGLYKLARTKIKGFTGIDDPYEPPIGEIVIMKDEECFSPKAMAKOVLCYLE 336
Db 212 CEARDPKGLYKLARAKIKGFTGIDDPYEPPLNCE--ISLREGGTSPTEMAEKVVGVYID 269
QY 337 ENGYLQA 343
Db 270 NKGYLQA 276

RESULT 13
AAG11015
ID AAG11015 standard; protein; 276 AA.
AC AAG11015;
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9571.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
```



```
XX PS Claim 1; SEQ ID NO 2270; 134pp; English.
XX CC This invention relates to a novel method for altering one or more plant
XX CC characteristics. Specifically, it refers to identifying genes that are up
XX CC - or down-regulated in transgenic plants overexpressing the heterodimeric
XX CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
XX CC alter plant characteristics accordingly. The present invention describes
XX CC generating transgenic plants for the production of growth regulators,
XX CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
XX CC the altered plant characteristics are selected from increased yield or
XX CC biomass, enhanced survival capacity, stress tolerance, plant architecture
XX CC or physiology, altered endoreduplication, biochemistry, signal
XX CC transduction, storage lipid mobilisation and/or altered photosynthesis,
XX CC each relative to the corresponding wild type plants. Accordingly, these
XX CC sequences can also be useful as positive or negative selectable markers
XX CC during transformation of cells or tissues. The identified genes play a
XX CC role in a variety of biological processes such as DNA replication, cell
XX CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
XX CC transcription factors. This polypeptide sequence is thale cress protein
XX CC expressed by a gene repressed 1.3 fold or more in plants overexpressing
XX CC the E2Fa/Dpa transcription factor, given in an exemplification of the
XX CC invention.
XX SQ Sequence 276 AA;
Query Match 44.7%; Score 816; DB 8; Length 276;
Best Local Similarity 64.4%; Pred. No. 1e-71;
Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;
QY 100 GARTCHRGIGRWRRRRNGAAGEAP---HSPVKEKPVMSNICKSTNIMHNLIGQS 156
Db 43 GSQTLSH-----NKGSIPEVKSINGHTGQKGP-LSTGVGNSINIKWHECSVEKV 91
QY 157 DRQKLLGQKGVVWTGLSGSKSTLACSLRELHCRGHLTYVLGDNLRHGLNRDLSEK 216
Db 92 DRQRLDQKGVVWTGLSGSKSTLACALNQLYQKGLCYLVDGDNVRHGLNRDLSEK 151
QY 217 AEDRAENIRRVGEVAKLPADAGVICIASLISPYRRDRDCAALLPHSNFIEVFDLPKI 276
Db 152 AEDRAENIRRVGEVAKLPADAGIICIASLISPYRTDRDCAKRLLEGGDFVEFMDVPLSV 211
QY 277 CEARDPKGLYKLARTGKIKGFTGIDDPYEPPIGEIVIMKDECECPSPKAMAKQVLCYLE 336
Db 212 CEARDPKGLYKLARAGKIKGFTGIDDPYEPPLNCE--ISLREGGCTSPTEMAEKVVGYLD 269
QY 337 ENGYLQA 343
Db 270 NRGYLQA 276
RESULT 15
AAG14478
ID AAG14478 standard; protein; 236 AA.
XX AC AAG14478;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 14356.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2005, 12:44:50 ; Search time 53 seconds
(without alignments)
483.106 Million cell updates/sec

Title: US-10-829-432-4
Perfect score: 1826
Sequence: 1 RPFHINQTEPLVTHQPP.....PKAMAKVLCYLENGYLQA 343

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	521.5	28.6	204	3	US-09-134-001C-4517
2	517	28.3	226	4	US-09-543-681A-7167
3	503	27.5	216	4	US-09-489-039A-9974
4	502.5	27.5	610	2	US-08-879-561-10
5	495	27.1	624	2	US-08-879-561-3
6	495	27.1	624	2	US-08-879-561-9
7	495	27.1	624	4	US-09-949-016-7339
8	494.5	27.1	641	4	US-09-949-016-7796
9	484.5	26.5	617	4	US-09-898-165B-8
10	482	26.4	615	4	US-09-898-165B-7
11	474	26.0	202	4	US-09-538-092-479
12	474	26.0	619	4	US-09-786-240-6
13	302	16.5	652	4	US-09-252-991A-24554
14	272.5	14.9	152	4	US-09-902-540-11522
15	222.5	12.2	244	4	US-09-902-540-15303
16	150.5	8.2	160	4	US-09-266-965-136
17	122	6.7	1039	4	US-09-252-991A-28966
18	119	6.5	656	4	US-09-252-991A-22270
19	108.5	5.9	1706	4	US-09-252-991A-21760
20	108	5.9	253	4	US-09-252-991A-29632
21	108	5.9	304	4	US-09-252-991A-16945
22	107.5	5.9	516	4	US-09-291-170A-1
23	107.5	5.9	516	4	US-09-724-884-1
24	107.5	5.9	808	4	US-09-252-991A-32826
25	106.5	5.8	210	4	US-09-252-991A-16748
26	106	5.8	388	4	US-09-252-991A-21572
27	105.5	5.8	494	4	US-09-252-991A-28683

28	103.5	5.7	262	4	US-09-252-991A-19030	Sequence 19030, A
29	103.5	5.7	666	4	US-09-050-739-70	Sequence 70, Appl
30	103	5.6	174	4	US-09-149-476-635	Sequence 635, App
31	103	5.6	242	4	US-09-252-991A-31425	Sequence 31425, A
32	103	5.6	298	4	US-09-252-991A-19524	Sequence 19524, A
33	103	5.6	472	4	US-09-252-991A-20457	Sequence 20457, A
34	103	5.6	1049	4	US-09-252-991A-25336	Sequence 25336, A
35	102.5	5.6	234	4	US-09-489-039A-12555	Sequence 12555, A
36	102	5.6	517	4	US-09-724-592-1	Sequence 1, Appli
37	102	5.6	527	4	US-09-252-991A-31093	Sequence 31093, A
38	102	5.6	585	4	US-09-252-991A-25568	Sequence 25568, A
39	102	5.6	599	4	US-09-252-991A-17250	Sequence 17250, A
40	101.5	5.6	313	4	US-09-252-991A-29545	Sequence 29545, A
41	101	5.5	276	4	US-09-252-991A-31870	Sequence 31870, A
42	101	5.5	971	4	US-09-538-092-1332	Sequence 1332, Ap
43	100.5	5.5	417	4	US-09-252-991A-18550	Sequence 18550, A
44	100.5	5.5	515	4	US-09-252-991A-32630	Sequence 32630, A
45	100	5.5	448	4	US-09-252-991A-24264	Sequence 24264, A

ALIGNMENTS

RESULT 1

US-09-134-001C-4517
; Sequence 4517, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4517
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4517

Query Match	28.6%	Score	521.5	DB	3	Length	204
Best Local Similarity	50.5%	Pred. No.	4.8e-45				
Matches	102	Conservative	34	Mismatches	63	Indels	3
						Gaps	1
QY	140	IGKSTNIIWHNCLIGQSDROKLLGQKGVVWITGLSGSGKSTLACALSRHLCHRLTVV	199				
Db	6	MSESNHITWHDSEVTKQROHKNHKSXVIFTGLSGSGKSTVSVLEKELFNEGKQTYR	65				
QY	200	LDGDNLRHGLNRDLSFKAEADRAENIRRVGEVAKLFADAGVTCIASLISPYRRDRDACRAL	259				
Db	66	LDGDNVRHGLNKLNGFSPEDRSNIRRIEVEVAKLMDAGALTVTAFISPYKEDREGVRL	125				
QY	260	LPHSNFIEVFDLPLKICEARDPKGLYKLTARTKIGFTGIDDDPYEPPIINGEIVIMKDE	319				
Db	126	LEDNEFIEVYTKSVECEKRDPKGLYKARSIGEFTGISAPYQAPENPEITI---DT	182				
QY	320	ECPSPKAMAKOVLCYLENGYL	341				
Db	183	EHDITQSVVQIIIRYLKEHYI	204				

RESULT 2

US-09-543-681A-7167
; Sequence 7167, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

```

; TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE:  2709.1002-001
; CURRENT APPLICATION NUMBER:  US/09/543.681A
; CURRENT FILING DATE:  2000-04-05
; PRIOR APPLICATION NUMBER:  US 60/128,706
; PRIOR FILING DATE:  1999-04-09
; NUMBER OF SEQ ID NOS:  8344
; SEQ ID NO 7167
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7167

Query Match      28.3%; Score 517; DB 4; Length 226;
Best Local Similarity 47.1%; Pred.No.1.6e-44;
Matches 107; Conservative 31; Mismatches 57; Indels 32; Gaps 4;

QY      133  EKPVMSNIGK-----TNLWHNCLIGSDRQKLLGQKGVVWITGLSGSGKST 181
DB      5  DPPFFALGRKRLTGRXIVIHQIDVWHVHQIGLKEREAQVHGKGVLWFTGLSGSGKST 64

QY      182  LACALGRELHCRGHL-----TYVL DGNLRHGLNRDLSFKAEDRAENIRRVGEVAKLPAD 236
DB      65  LADALEQTLQYSTLHAPIRTYLLDGNLRHGLCHDLGFSQDRHENIRRVGEVAKLMVD 124

QY      237  AGVICTASLISPYRRDRDACEALLPHSNFEVFTDLPLKICEARDPGLYKLARTGKIG 296
DB      125  AGLIVLTAFTISPYQQDRQVRERPAQGRFIEIFVDTPLALCEARDPKGLYQKARRGEIKQ 184

QY      297  FTGIDDPYEP-----INGEIVIKMKDECPSPKAMQVLCYLEE 337
DB      185  FSGIDSPYEPTAPEIHLGDSLAI-----NELTQQILAVLQQ 221

RESULT 3
US-09-489-039A-9974
; Sequence 9974, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE:  2709.2004001
; CURRENT APPLICATION NUMBER:  US/09/489,039A
; CURRENT FILING DATE:  2000-01-27
; PRIOR APPLICATION NUMBER:  US 60/117,747
; PRIOR FILING DATE:  1999-01-29
; NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 9974
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9974

Query Match      27.5%; Score 503; DB 4; Length 216;
Best Local Similarity 57.7%; Pred.No.4e-43;
Matches 97; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

QY      145  NILWHNCLIGSDRQKLLGQKGVVWITGLSGSGKSTLACALSRELHCRGHLTYVL DGN 204
DB      22  NVVWHAHPVTQQREQHGHGCVVWITGLSGSGKSVAGALEALHERGVSTYLL DGN 81

QY      205  LRHGLNRDLSFKAEDRAENIRRVGEVAKLPADAGVICIASLISPYRRDRDACEALLPHSN 264
DB      82  VRHGLCSDLGFSDBDRKENIRRVGEVARLWMDAGLVLTAFISPHRAERQWVRERLGEGR 141

QY      265  FIEVFIDPLKICEARDPGLYKLARTCKIKGFTGIDDPYEPPIGEI 312
DB      142  FIEVFVDTPLAICEARDPGLYKKARAGELRNFTGIDSVYEAPKA EI 189

RESULT 4
US-08-879-561-10

```



```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7339
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-7339

Query Match 27.1%; Score 495; DB 4; Length 624;
Best Local Similarity 46.1%; Pred. No. 1.2e-41;
Matches 106; Conservative 38; Mismatches 60; Indels 26; Gaps 6;

Qy 125 EAPHSPVKPKVMN-----IKSTNLIWHNCLIGQSDRQKLLGQ----KGVVWITGLS 175
Db 2 EIPGSLCKVKLSNNAQMGORATNVTYQAHVSRNKRQGVVGRGFRGCTVWLTGLS 61

Qy 176 GSKSTLACALSRHLHCRGHLTVVLGDGNLRHGLNRLDSFKAEDRAENIRRVGEVAKLFA 235
Db 62 GAGKTTVWALREYLVCHGIPCYTLDGNIRQLGNLKNLGFSPEDREENVRRIAEVAKLFA 121

Qy 236 DAGVICIASLISPYRRDRDADR-----ALLPHSNFIEVFDLPKICBARDPKGLYKLAR 290
Db 122 DAGLVCTISFISPTQDRNNAQIHGASLP---FFEIVFDAPLVHCEQRDVKGLYKAR 178

Qy 291 TGKIKFTGIDDPYEPPIGEIVIKWKD---RECSPKAMAKOVLVLEE 337
Db 179 AGEIKGFTGIDSEYKPEAPELVLTDCDVNDC-----VQVVVELLQOE 222

RESULT 8
US-09-949-016-7796
; Sequence 7996, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7796
; LENGTH: 641
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-7796

Query Match 27.1%; Score 494.5; DB 4; Length 641;
Best Local Similarity 49.3%; Pred. No. 1.4e-41;
Matches 105; Conservative 32; Mismatches 59; Indels 17; Gaps 5;

Qy 117 RRNGAARG-EAPHSPVK---BKPVMSNTGKSTNLIWHNCLIGQSDRQKLLGQ----KGC 167
Db 11 RPSVSPAGTRAPPQPPASMSGIKKQKTNQVQAHVSRNKRQGVVGRGFRGC 70

Qy 168 VVWITGLSGSKSTLACALSRHLHCRGHLTVVLGDGNLRHGLNRLDSFKAEDRAENIRRV 227

```

```

RESULT 13
US-/09-252-9154A
; Sequence 24554, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AL
; TITLE OF INVENTION: AERUGINOSA FOR DIA
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,91
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,78
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,19
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

```

```
; SEQ ID NO 24554
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24554

Query Match      16.5%; Score 302; DB 4; Length 652;
Best Local Similarity 29.7%; Pred. No. 6.4e-22;
Matches 94; Conservative 37; Mismatches 113; Indels 72; Gaps 9;

QY 52 QRAPVPLGTLTSDAPLPAIVHGL-----TPRSS-----HSSAGIASDSGR 93
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 375 KRATSVPGSIPS-----IVHKVDVNTLRTPGSELKLNELARVKVSLDAPIALD-GY 426
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 94 REGEGRGARTHCHRGIVRRRENG-----AAPGEAPHSPVKEKPMWSNIG 141
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 427 EQNRITGAPI-----VIDRLTNGTVGAGMIISAPSAHSSAHH----- 465
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 142 KSTNLTWNCILGSDRQKLLQKQGVVWITGLSGSKSTLACALSRELHCRGHLTYVLD 201
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 466 -----GSAHVTRERACRFQOQATVLFSLGSGAGKSTLAYAVERKLFDMGRAVYVLD 519
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 202 GNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACLALP 261
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 520 GQNLRHDLNKLGPQDRAGTGNWLRTHAVAKQFNEAGLISLCFAVAPSAGREGQARALIG 579
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 262 HSNFIEVFIDLPLKICEARDPKGLYKLAFTGKIKGFTGIDDPYEPPINGEIVIKMDEEC 321
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 580 AERLITVVVQASPOVCRENDPOGLYAAGE-----DNIPGESFYDVPDLADLVI----DTQA 632
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 322 PSPKAMAKQVLCYLEE 337
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 633 LSVEDGVKQVLDLRE 648
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

RESULT 14
US-09-902-540-11522
; Sequence 11522, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 11522
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-11522

Query Match      14.9%; Score 272.5; DB 4; Length 152;
Best Local Similarity 42.2%; Pred. No. 8.4e-20;
Matches 62; Conservative 23; Mismatches 57; Indels 5; Gaps 2;

QY 199 VLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDADRA 258
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 5 VLDGDEVRTLWLRGLGFTREDRENVRRIGHVARLLAKHGVGVIAAISPYASSRAEVR 64
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 259 LLPHSN--FIEVFIDLPLKICEARDPKGLYKLAFTGKIKGFTGIDDPYEPPINGEIVIKM 316
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 65 LAEEAGIPFEIVVQAPLDVLIARDVKGLYKKALAGLAHFTGVSDDPYEAPDAPDVTV-- 122
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 317 KDEECPSPKAMAKQVLCYLEENGYLQA 343
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 123 -HSDVDVTVEAGLWRVLETLRKRLDLDA 148
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
```

RESULT 15

```
US-09-902-540-15303
; Sequence 15303, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15303
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15303
```

```
Query Match      12.2%; Score 222.5; DB 4; Length 244;
Best Local Similarity 34.5%; Pred. No. 2.1e-14;
Matches 58; Conservative 23; Mismatches 82; Indels 5; Gaps 2;
```

```
QY 174 LSGSGKSTLACALSRELHCRGHLTYVLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKL 233
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 1 MSGTGKSTTAAITAAARLRQVGRNVVEILDEGELGELAGLWAGLDGDKEDRITTVVRLGJVANL 60
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 234 FADAGVICIASLISPYRRDRDACLALPHSNFIEVFIDLPLKICEARDPKGLYKLAFTGK 293
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 61 LTRNGVAALVPCVSPYKPGREENRSI--GRYVEVVYDCPTKLIERDSTGRYKKALNGE 118
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 294 IKGFTGIDDPYEPPINGEIVIKMKDEECPSPKAMAKQVLCYLEENGYL 341
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 119 IPNFIGITEPYEPPNSPEVTIHSNVESVEDGAAKIFQSLDL---GYM 163
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
```

Search completed: September 9, 2005, 13:05:05
Job time : 53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2005, 12:58:46 ; Search time 112 Seconds
(without alignments)
1207.946 Million cell updates/sec

Title: US-10-829-432-4
Perfect score: 1826
Sequence: 1 RPFHFNTEPLVTHTQPP.....PKMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1777461 seqs, 394431504 residues

Total number of hits satisfying chosen parameters: 1777461

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1826	100.0	343	16	US-10-829-432-4
2	1168.5	64.0	270	16	US-10-425-115-301616
3	1013.5	55.5	253	15	US-10-425-114-47513
4	972	53.2	207	16	US-10-739-930-10868
5	936	51.3	188	16	US-10-767-701-40866
6	903	49.5	429	16	US-10-425-115-247552
7	900	49.3	304	16	US-10-437-963-154262
8	891.5	48.8	274	15	US-10-425-114-41817
9	887	48.6	224	16	US-10-829-432-2
10	857	47.5	349	16	US-10-425-115-319790
11	858	47.0	246	16	US-10-829-432-10
					Sequence 4, Appli
					Sequence 301616, A
					Sequence 47513, A
					Sequence 10868, A
					Sequence 40866, A
					Sequence 247552,
					Sequence 154262,
					Sequence 41817, A
					Sequence 2, Appli
					Sequence 319790,
					Sequence 10, Appl

12	854	46.8	215	15	US-10-425-114-55578	Sequence 55578, A
13	854	46.8	265	15	US-10-424-599-205747	Sequence 205747,
14	853	46.7	207	15	US-10-424-599-262462	Sequence 262462,
15	852	46.7	311	16	US-10-829-432-8	Sequence 8, Appli
16	845	46.3	234	15	US-10-425-114-61644	Sequence 61644, A
17	839	45.9	246	15	US-10-425-114-47070	Sequence 47070, A
18	833	45.6	312	16	US-10-829-432-13	Sequence 13, Appl
19	830	45.5	246	15	US-10-425-114-66421	Sequence 66421, A
20	816	44.7	276	16	US-10-829-432-14	Sequence 14, Appl
21	777.5	42.6	293	14	US-10-342-224-66	Sequence 66, Appl
22	770	42.2	173	16	US-10-437-963-175387	Sequence 175387,
23	723.5	39.6	301	16	US-10-425-115-219699	Sequence 219699,
24	699	38.3	207	15	US-10-424-599-262463	Sequence 262463,
25	674.5	36.9	185	16	US-10-425-115-301613	Sequence 301613,
26	611	33.5	158	16	US-10-437-963-148349	Sequence 148349,
27	601.5	32.9	203	16	US-10-425-115-219698	Sequence 219698,
28	571	31.3	200	15	US-10-282-122A-51644	Sequence 51644, A
29	555	30.4	199	15	US-10-369-493-17684	Sequence 17684, A
30	534.5	29.3	202	15	US-10-369-493-17218	Sequence 17218, A
31	529.5	29.0	162	9	US-09-795-693-40	Sequence 40, Appl
32	529.5	29.0	162	14	US-10-156-239-40	Sequence 40, Appl
33	529.5	29.0	162	14	US-10-199-485-40	Sequence 40, Appl
34	521.5	28.6	199	15	US-10-282-122A-70988	Sequence 70988, A
35	521.5	28.6	204	18	US-10-724-972A-5067	Sequence 5067, Ap
36	519.5	28.5	196	15	US-10-369-493-16684	Sequence 16684, A
37	514.5	28.2	197	15	US-10-282-122A-46511	Sequence 46511, A
38	514.5	28.2	204	15	US-10-282-122A-68949	Sequence 68949, A
39	510	27.9	213	15	US-10-282-122A-77726	Sequence 77726, A
40	505.5	27.7	199	15	US-10-282-122A-71917	Sequence 71917, A
41	505	27.7	652	15	US-10-369-493-6124	Sequence 6124, Ap
42	503	27.5	201	15	US-10-282-122A-59084	Sequence 59084, A
43	502	27.5	127	16	US-10-767-701-61157	Sequence 61157, A
44	498.5	27.3	660	15	US-10-369-493-17600	Sequence 17600, A
45	496.5	27.2	621	15	US-10-369-493-10735	Sequence 10735, A

ALIGNMENTS

RESULT 1
US-10-829-432-4
; Sequence 4, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Zea mays
US-10-829-432-4

Query Match	100.0%	Score	1826;	DB	16;	Length	343;
Best Local Similarity	100.0%	Pred. No.	9.5e-154;				
Matches	343;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Oy	1	RPFHFNTEPLVTHTQPPSPAPASQOGQNGNTLLSPPTTLAVILVNPQRAPPVLP	60				
Db	1	RPFHFNTEPLVTHTQPPSPAPASQOGQNGNTLLSPPTTLAVILVNPQRAPPVLP	60				
Oy	61	LTPSDAPLPAIVHGLTPRSSHSSAGLASDSGRREGRGARTCHRGIGRWVRRRRNG	120				

Db 61 LTPSDAPLALVIHGLTPRSHSSAGLASDSGRREGGARGARTHCHRGIGRWVRRRRNG 120
Qy 121 AAPGEAPHSPVKEKPVMSNIGKSTNIIWHNCLIGSDRQKLLGQKGCVVWITGLSGSGKS 180
Db 121 AAPGEAPHSPVKEKPVMSNIGKSTNIIWHNCLIGSDRQKLLGQKGCVVWITGLSGSGKS 180
Qy 181 TLACALSRELHCRGHLTYVLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGVI 240
Db 181 TLACALSRELHCRGHLTYVLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGVI 240
Qy 241 CIASLISPYRRDRDACRALLPHSNFIEVFDLPLKICEARDPKGLYKLARTGKIKGFTGI 300
Db 241 CIASLISPYRRDRDACRALLPHSNFIEVFDLPLKICEARDPKGLYKLARTGKIKGFTGI 300
Qy 301 DDPEYPPINGEIVIKKDEECSPKAMAKQVLCYLEENGYLQA 343
Db 301 DDPEYPPINGEIVIKKDEECSPKAMAKQVLCYLEENGYLQA 343

RESULT 2

US-10-425-115-301616
; Sequence 301616, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 301616
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: M74577_38151C.1.ppep
US-10-425-115-301616

Query Match 64.0%; Score 1168.5; DB 16; Length 270;
Best Local Similarity 84.5%; Pred. No. 2.1e-95;
Matches 235; Conservative 6; Mismatches 28; Indels 9; Gaps 4;

Qy 66 APLPALVIHGLTPRSHSSAGLASDSGRREGGARGARTHCHRGIGRWVRRRRNGAAPGE 125
Db 2 ASLP--VPHTL--PRASPAIVGAA--RGRAAVRVRTATATGALGGCGGGGMEQRH---GE 52
Qy 126 APHPVKEKPVMSNIGKSTNIIWHNCLIGSDRQKLLGQKGCVVWITGLSGSGKSTLACA 185
Db 53 APHPVKEKPVMSNIGKSTNIIWHNCLIGSDRQKLLGQKGCVVWITGLSGSGKSTLACA 112
Qy 186 LSRELHCRGHLTYVLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGVICIASL 245
Db 113 LSRELHCRGHLTYVLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGVICIASL 172
Qy 246 ISPYRRDRDACRALLPHSNFIEVFDLPLKICEARDPKGLYKLARTGKIKGFTGIDDPYE 305
Db 173 LSPYRRDRDACRALLPHSNFIEVFDLPLKICEARDPKGLYKLARTGKIKGFTGIDDPYE 232
Qy 306 PPINGEIVIKKDEECSPKAMAKQVLCYLEENGYLQA 343
Db 233 PPINGEIVIKKDEECSPKAMAKQVLCYLEENGYLQA 270

RESULT 3

US-10-425-114-47513
; Sequence 47513, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47513
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700097023_FLI.ppep
US-10-425-114-47513

Query Match 55.5%; Score 1013.5; DB 15; Length 253;
Best Local Similarity 79.8%; Pred. No. 1.2e-81;
Matches 210; Conservative 4; Mismatches 20; Indels 29; Gaps 6;

Qy 64 SDAPLPALVIHGLTPRSHSSAGLASDSGRREGGARGA---RT-----HCHRGIGRWV 113
Db 1 SMASLP--VPHTL--PRASPAIVGAA-----RGRAAVRVRTATATGALGGCGGGGMEQ 49
Qy 114 RRRRNGAACEAPHSPVKEKPVMSNIGKSTNIIWHNCLIGSDRQKLLGQKGCVVWITG 173
Db 50 RH-----GEAPHSPVKEKPVMSNIGKSTNIIWHNCLIGSDRQKLLGQKGCVVWITG 101
Qy 174 LSGSGKSTLACALSRELHCRGHLTYVLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKL 233
Db 102 LSGSGKSTLACALSRELHCRGHLTYVLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKL 161
Qy 234 PADAGVICIASLISPYRRDRDACRALLPHSNFIEVFDLPLKICEARDPKGLYKLARTGK 293
Db 162 FSDAGVICIASLISPYRRDRDACRALLPHSNFIEVFDLPLKICEARDPKGLYKLARTGK 221
Qy 294 IKGTGIDDPYEPPINGEIVIKM 316
Db 222 IKGTGIDDPYEPPINGEIVIKM 244

RESULT 4

US-10-739-930-10868
; Sequence 10868, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 10868
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-C7564_1.p
US-10-739-930-10868

Query Match 53.2%; Score 972; DB 16; Length 207;
Best Local Similarity 87.9%; Pred. No. 4.5e-78;
Matches 182; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

Qy 137 MSNIGKSTNIIWHNCLIGSDRQKLLGQKGCVVWITGLSGSGKSTLACALSRELHCRGHL 196
Db 1 MSNIGKSTNIIWHNCLIGSDRQKLLGQKGCVVWITGLSGSGKSTLACALSRELHCRGHL 60
Qy 197 TVVLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDAC 256

Db 61 TVVLDGDNLRHGLNLSFKAEDRTENIRRVGEVAKLFADAGTICIASLISPYRRDRDAC 120
QY 257 RALLPHSNFIEVFDLPLKICARPKGLYKLTARTKGTGIDDPYEPPIRVIVKM 316
Db 121 RALLPDSRFIEVMDLPLELCSARPDKGLYKLTARTKGTGVDPPYSPVNSEIVKM 180
QY 317 KDEECPSPKAMAKQVLCYLENGYLQA 343
Db 181 EGGECPSPKAMAKQVLCYLENGYLQA 207

RESULT 5
US-10-767-701-40866
; Sequence 40866, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; NUMBER OF SEQ ID NOS: 63128
; CURRENT FILING DATE: 2004-01-29
; SEQ ID NO 40866
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C114933_1.pap
US-10-767-701-40866

Query Match 51.3%; Score 936; DB 16; Length 188;
Best Local Similarity 94.7%; Pred. No. 6.4e-75;
Matches 178; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 156 SDRQKLGQKGCVMWITGLSGSGKSTLACALSRELHCRGHLTYVLDGDNLRHGLNRLDSF 215
Db 1 SDRQKLGQKGCVMWITGLSGSGKSTLACALSRELHCRGHLTYVLDGDNLRHGLNRLDSF 60
QY 216 KAEDRAENIRRVGEVAKLFADAGVCIASLISPYRRDRDACLALPHSNFIEVFDLPLK 275
Db 61 KAEDRAENIRRVGEVAKLFADAGVCIASLISPYRRDRDACLALPHSNFIEVFDLPLK 120
QY 276 ICEARDPKGLYKLTARTKGTGIDDPYEPPIRINGEIVIMKDECPSPKAMAKQVLCYL 335
Db 121 ICEARDPKGLYKLTARTKGTGIDDPYEPPIRINGEIVIMKDECPSPKAMAKQVLCYL 180
QY 336 BENGYLQA 343
Db 181 BENGYLQA 188

RESULT 6
US-10-425-115-247552
; Sequence 247552, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 247552
; LENGTH: 429
; TYPE: PRT

; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(429)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_157357C.1.pap
US-10-425-115-247552
Query Match 49.5%; Score 903; DB 16; Length 429;
Best Local Similarity 53.7%; Pred. No. 1.6e-71;
Matches 189; Conservative 33; Mismatches 98; Indels 32; Gaps 4;
QY 7 NOTEPLVTHTOQPP-----SPAGPASQOQOGNTLL-----SPTPTTLAVILV 49
Db 91 SRTPRTAHGRTPPCALVLSVSRLPDPGARGGGALAILVSTIPAAGVFPFPLAVARA 150
QY 50 NPQRAPPVLPGLTPSDAPLPALVIHGLTPRSHSASGLASDSGRREGSGRGARTHCHRG 109
Db 151 SRXRA-----PEPEPRRESRVASASPPPPPETPSGRNXDEGNRRQRLXGRPLVEGAG 202
QY 110 GRWVRRRRNGAACEAPHSPVKEKPVMSNIGKSTNLIWHNCLIGOSDROKLGKGCVV 169
Db 203 DRSVEQEPHAGVEGQA-----LXMSSTVPKSSNIFMHDCLVGTQKTDQKLLNQKCV 255
QY 170 WITGLSGSGKSTLACALSRELHCRGHLTYVLDGDNLRHGLNRLDSFKAEDRAENIRRVGE 229
Db 256 WITGLSGSGKSTLACTLGLRELHTRGKLAYVLDGDNLRHGLNKLDFKAEDRAENIRRVGE 315
QY 230 VAKLFADAGVCIASLISPYRRDRDACLALPHSNFIEVFDLPLKICEARDPKGLYKLA 289
Db 316 VAKLFADAGVCIASLISPYRRDRDACLALPHSNFIEVFDLPLKICEARDPKGLYKLA 375
QY 290 RTGKIKGTGIDDPYEPPIRINGEIVIMKDECPSPKAMAKQVLCYLENGYL 341
Db 376 RAGKIKGTGIDDPYEPPIRINGEIVIMKDECPSPKAMAKQVLCYLENGYL 427

RESULT 7
US-10-437-963-154262
; Sequence 154262, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 154262
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5413C.1.pap
US-10-437-963-154262
Query Match 49.3%; Score 900; DB 16; Length 304;
Best Local Similarity 57.0%; Pred. No. 1.9e-71;
Matches 188; Conservative 27; Mismatches 61; Indels 54; Gaps 4;
QY 23 APGASQOQOGNTLLSPTTLA--VILVNPQRPVPLPGLTPSDAPLPALVHGLTPRS 80
Db 18 APAAAADGRLAGGVLRPPWGGVAAAGVRLVAA--PMS 53

```
QY 81 SHSSAGLASDSRRREGGEGRGARTHCHRGIGRWVRRRRRNGAAPGEAPHSS-----PVK 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 54 VVTAAGSGSWDRKAG-----RSEAAVPAAPVDGKDDDLVL 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 133 EKPVMNIGKSTNIIWHNCLIGSDRQKLLGQGVVITGLSGSKSTLACALSREIHC 192
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 94 NKLMTSTVGKSTNIRWHDPCVNLQDRQKLLNGQGVVITGLSGSKSTLACALSREIHS 153
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 193 RGHLYVLVDGDNLRHGLNRLSFKAEADRAENIRRVGEVAKLPADAGVICIASLISPYRRD 252
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 154 ROHLTYVLVDGDNLRHGLNRLSFKAKDRAENIRRVGEVAKLPADAGLICITSLISPYKSD 213
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 253 RDCRALLPHSNFIEVFDLPKICEARDPKGLYKLARTGKIKGFTGIDDDPYEPPINGEI 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 214 RSACKLLPNSSFIEVFLNVLPEVCEERDPKGLYKLARAGKIKGFTGIDDDPYETPSDCEI 273
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 313 VTKMDCECPSPKAMAKQVLCYLENGYLQ 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 274 VIQCKVGDCPSKSMADQVVSYLEANGPFQ 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
US-10-425-114-41817
; Sequence 41817, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 41817
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3066-014-B1_FLI.pep
US-10-425-114-41817

Query Match 48.8%; Score 891.5; DB 15; Length 274;
Best Local Similarity 65.8%; Pred. No. 9.5e-71;
Matches 183; Conservative 20; Mismatches 66; Indels 9; Gaps 3;

QY 66 APLPALVTHGLTPRSSHSSAGLASDSGRREGGEGRGARTHCHRGIGRWVRRRRRNG-AAPG 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 4 APPRPCCSGVCIAAHPPAAVA---ARPGTTRTTT-----VAAAAAASNGSAAAA 55
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 125 EAPHSPVKEKPVMSNIGKSTNIIWHNCLIGSDRQKLLGQGVVITGLSGSKSTLAC 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 56 VAGISSSSSALVTSTVGKSTNIIWHNCLIGSDRQKLLGQGVVITGLSGSKSTLAC 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 185 ALSRELHCRGHLYTYVLVDGDNLRHGLNRLSFKAEADRAENIRRVGEVAKLPADAGVICIAS 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 116 ALSRELHCRGHLYTYVLVDGDNLRHGLNRLSFKAEADRAENIRRVGEVAKLPADAGVICIAS 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 245 LISPYRRDRDACRALLPHSNFIEVFDLPKICEARDPKGLYKLARTGKIKGFTGIDDPY 304
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 176 LISPYRRDRSACRDLPLPKHSFIEVFDLPVLPQVCEARDPKGLYKLARAGKIKGFTGIDDPY 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 305 EPPINGEIVIKMKDECPSPKAMAKQVLCYLENGYLQ 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 EPPSDCEIIVIOCKVGDCPSPSMAGHVVSYLETNGFLQ 273
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-10-829-432-2

Query Match 47.5%; Score 867; DB 16; Length 349;
Best Local Similarity 54.0%; Pred. No. 2e-68;
Matches 190; Conservative 31; Mismatches 93; Indels 38; Gaps 6;
```

```
; Sequence 2, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Zea mays
US-10-829-432-2

Query Match 48.6%; Score 887; DB 16; Length 224;
Best Local Similarity 77.0%; Pred. No. 1.8e-70;
Matches 171; Conservative 16; Mismatches 35; Indels 0; Gaps 0;

QY 121 AAPGEAPHSPVKEKPVMSNIGKSTNIIWHNCLIGSDRQKLLGQGVVITGLSGSKS 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 AAAAVAGITSSSSSALVTSTVGKSTNIIWHNCLIGSDRQKLLGQGVVITGLSGSKS 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 TLACALSRELHCRGHLYTYVLVDGDNLRHGLNRLSFKAEADRAENIRRVGEVAKLPADAGVI 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 TLACALSRELHCRGHLYTYVLVDGDNLRHGLNRLSFKAEADRAENIRRVGEVAKLPADAGLV 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 CIASLISPYRRDRDACRALLPHSNFIEVFDLPKICEARDPKGLYKLARTGKIKGFTGI 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 CIASLISPYRRDRSACRDLPLPKHSFIEVFDLPVLPQVCEARDPKGLYKLARAGKIKGFTGI 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 301 DDVPEPPINGEIVIKMKDECPSPKAMAKQVLCYLENGYLQ 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 DDVPEPPSDCEIIVIOCKVGDCPSPSMAGHVVSYLETNGFLQ 223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-10-425-115-319790
; Sequence 319790, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 319790
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(349)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_54717C.1.pep
US-10-425-115-319790

Query Match 47.5%; Score 867; DB 16; Length 349;
Best Local Similarity 54.0%; Pred. No. 2e-68;
Matches 190; Conservative 31; Mismatches 93; Indels 38; Gaps 6;
```


Db 59 MSNVANSTNIMWHDPIQKDRQQLQKQKGVWITGLSGSGKSTIACALSRSLHSGKL 118
Qy 197 TVVLGDGNLRHGLNRLDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDAC 256
Db 119 SYVLGDGNIRHGLNQDLSFRAEDRSNIRRIIGEVAKLFDADAGVICITSLISPYQKDRDAC 178
Qy 257 RALLPHSNFIEVFIDPLKICEARDPKGLYKLARTGKIKGFTGIDDPYEPPIINGEIVIKM 316
Db 179 RALIPKGFIEVFIDVPLHVCEARDPKGLYKLARAGKIKSFTGIDDPYEPPIINGEIVIKM 238
Qy 317 KDEECPSPKAMAKQVLCYLENGYLQA 343
Db 239 KGSCKSPSDMAEEVISYLENGYLRA 265

RESULT 14
US-10-424-599-262462
; Sequence 262462, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 262462
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_79025C.1.pap
US-10-424-599-262462

Query Match 46.7%; Score 853; DB 15; Length 207;
Best Local Similarity 75.4%; Pred. No. 1.7e-67;
Matches 156; Conservative 24; Mismatches 27; Indels 0; Gaps 0;
Qy 137 MSNIGKSTNIMWHDPIQKDRQQLQKQKGVWITGLSGSGKSTIACALSRSLHSGKL 196
Db 1 MATLSNSTNIFWQDCQIGRPERQKLLNQKGVWITGLSGSGKSTIACALSRSLHSGKL 60
Qy 197 TVVLGDGNLRHGLNRLDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDAC 256
Db 61 SYVLGDGNLRHGLNQDLSFKAEDRTENIRRTIGEVAKLFDADAGLICVASLISPYRRDRDTC 120
Qy 257 RALLPHSNFIEVFIDPLKICEARDPKGLYKLARTGKIKGFTGIDDPYEPPIINGEIVIKM 316
Db 121 RAMLPDANFIEVFMMNPELCEARDPKGLYKLARAGKIKGFTGIDDPYEPPIINGEIVIKM 180
Qy 317 KDEECPSPKAMAKQVLCYLENGYLQA 343
Db 181 ENGNCPPTMTMAGQVVTYLENGGFLES 207

RESULT 15
US-10-829-432-8
; Sequence 8, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Glycine max
US-10-829-432-8
Query Match 46.7%; Score 852; DB 16; Length 311;
Best Local Similarity 73.9%; Pred. No. 3.6e-67;
Matches 153; Conservative 31; Mismatches 23; Indels 0; Gaps 0;
Qy 137 MSNIGKSTNIMWHDPIQKDRQQLQKQKGVWITGLSGSGKSTIACALSRSLHSGKL 196
Db 105 MSNVGNSTNIMWHDPIQKDRQQLQKQKGVWITGLSGSGKSTIACALSRSLHSGKL 164
Qy 197 TVVLGDGNLRHGLNRLDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDAC 256
Db 165 SYVLGDGNIRHGLNQDLSFRAEDRSNIRRIIGEVAKLFDADAGVICITSLISPYQKDRDAC 224
Qy 257 RALLPHSNFIEVFIDPLKICEARDPKGLYKLARTGKIKGFTGIDDPYEPPIINGEIVIKM 316
Db 225 RALLSKGDFIEVFIDVPLHVCEARDPKGLYKLARAGKIKGFTGIDDPYEPPIINGEIVIKM 284
Qy 317 KDEECPSPKAMAKQVLCYLENGYLQA 343
Db 285 KGSCKSPSDMAEEVISYLENGYLRA 311

Search completed: September 9, 2005, 13:07:10
Job time : 113 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 9, 2005, 13:03:11 ; Search time 4068 Seconds
(without alignments)
4085.580 Million cell updates/sec

Title: US-10-829-432-4

Perfect score: 1826

Sequence: 1 RPFHFINQTEPLVTHTOQPP.....PRMAKQVLCYLENGYLQA 343

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10829432/runat_08092005_132223_25075/app_query.fasta_1.519
-DB=GenEmbl -QFMT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10829432 @CGN_1_1_5600 @runat_08092005_132223_25075 -NCPU=6 -ICPU=3
-DEV MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-NO TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_rc.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	876	48.0	928	8	BT009518 Triticum
2	874.5	47.9	1434	8	AK071285 Oryza sat
3	833	45.6	1086	8	AF044285 Catharant
4	816	44.7	831	6	CQ805858 Sequence

5	816	44.7	831	6	AX412633 Sequence
6	816	44.7	831	6	AX506015 Sequence
7	816	44.7	831	6	AV132010 Arabidops
8	816	44.7	1068	8	AY085264 Arabidops
9	816	44.7	1077	8	ATU05238 Arabidops
10	816	44.7	1143	8	AY054287 Arabidops
11	816	44.7	1185	8	ATCLAFPSK A.thaliana
12	815.5	44.7	970	8	AF178976 Zea mays
13	778.5	42.6	920	6	BD248389 Gene part
14	778.5	42.6	1243	8	AF462823 Arabidops
15	778.5	42.6	1311	8	AF043351 Arabidops
16	777.5	42.6	882	8	AY097421 Arabidops
17	757.5	41.5	964	8	BT005193 Arabidops
18	757.5	41.5	1143	8	BT003977 Arabidops
19	752.5	41.2	1133	8	AY085031 Arabidops
20	664	36.4	133022	8	AC118346 Oryza sat
21	633.5	34.7	3489	8	ATU59759 Arabidops
22	633.5	34.7	89840	8	AC004705 Arabidops
23	632.5	34.6	51860	8	AB013390 Arabidops
24	629.5	34.5	112067	8	AL161596 Arabidops
25	629.5	34.5	118267	8	ATCHRIV92 Arabidops
26	599	32.8	110000	1	AT5J17 Arabidops
27	595.5	32.6	1017	8	AF006618_26 Arabidops
28	592.5	32.4	302325	1	AF069952 Enteromor
29	587.5	32.2	101410	8	AE017236 Mycobacte
30	577	31.6	11131	1	ATAC009540 Arabidops
31	572.5	31.4	58756	2	AE007523 Clostridi
32	565.5	31.0	53533	1	Continuation (5 of
33	565.5	31.0	198743	1	Y09476 Bacillus su
34	563	30.8	2372	3	Z99109 Bacillus su
35	563	30.8	2378	3	AK173431 Ciona int
36	560	30.7	192657	1	AB036852 Ciona int
37	556.5	30.5	110000	1	AP005377 Thermosyn
38	552	30.2	153950	8	Continuation (24 o
39	551	30.2	2380	3	AC104429 Oryza sat
40	549.5	30.1	300150	1	AK173392 Ciona int
41	549	30.1	147205	8	AP004598 Oceanobi
42	547	30.0	1045	8	AC146702 Ctenomac
43	542	29.7	302275	1	AV325892 Porphyra
44	538.5	29.5	302050	1	AE016751 Staphyloc
45	537.5	29.4	71349	8	AL935256 Lactobaci
					AP006662 Lotus cor

ALIGNMENTS

RESULT 1	BT009518	Triticum aestivum clone wrl.pk0101.e2:fis, full insert mRNA	928 bp	linear	PLN 20-JUN-2003
LOCUS	BT009518	Triticum aestivum clone wrl.pk0101.e2:fis, full insert mRNA	928 bp	linear	PLN 20-JUN-2003
DEFINITION	BT009518	Triticum aestivum clone wrl.pk0101.e2:fis, full insert mRNA	928 bp	linear	PLN 20-JUN-2003
ACCESSION	BT009518	Triticum aestivum (bread wheat) sequence.			
VERSION	BT009518.1	GI:32129069			
KEYWORDS	FLI CDNA.				
SOURCE	Triticum aestivum				
ORGANISM	Triticum aestivum				
REFERENCE	1	(bases 1 to 928)			
AUTHORS	Tingey,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H., Caraher,N.R., Hanafey,M.K. and Hainey,C.F.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA				
FEATURES	Location/Qualifiers				
source	1..928				
	/organism="Triticum aestivum"				
	/mol_type="mRNA"				
	/db_xref="taxon:4565"				
	/clone="wrl.pk0101.e2:fis"				

Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers
 1. .1434
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J023086D14"

ORIGIN

Alignment Scores:

Pred. No.: 2,23e-33 Length: 1434
 Score: 874.50 Matches: 191
 Percent Similarity: 63.74% Conservative: 34
 Best Local Similarity: 54.11% Mismatches: 68
 Query Match: 47.89% Indels: 62
 DB: 8 Gaps: 7

US-10-829-432-4 (1-343) x AK071285 (1-1434)

QY 49 ValAsnProGlnArgAlaProProValLeuProGlyLeuThrProSerAspAlaProLeu 68
 DB 6 GTTAAACCCATTCGGGGGCCCTCCATCGATGGAGCGCTCTCCCGCTTTCACACACCA 65
 QY 69 ProAlaLeuValIleHis--GlyLeuThr-----ProArgSerSerHisSer 84
 DB 66 CCCCCTGCTCTCCACCGCGCGCATCACCGCGCCGCCCTCACTCTCTCTCTCTCTCC 125
 QY 85 Ala-----GlyLeuAlaSerAspSerGlyArgGluGlyGly 98
 DB 126 GCGAGACCGCGTGCACCTGCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 185
 QY 99 ArgGlyAlaArgThr-----GlyLeuAlaSerAspSerGlyArgGluGlyGly 103
 DB 186 GAGCCCGGCAATCTAGGTTGGCCGCCAATCTCCCGACGCTCCCGCTCCGCTCGC 245
 QY 104 -----HisCysHisArgGlyIleGlyArg----- 111
 DB 246 GCGCGGAGGATAACCGTGCAGTCACGGAG-GGCGCGTGGTGGTGGTGGTGGTGGTGG 304
 QY 112 -----Trp----- 112
 DB 305 CGCCTCTCTGGATCG 364
 QY 113 ---ValArgArgArgArgArgAlaAlaProGlyGluAlaProHisSerProVal 131
 DB 365 AGGAGGAGGAGGAGGAGGAGGAGCGC-TCTTCAGCGCGCGCGCGCGCGCGCGCGCG 423
 QY 132 LysGluLysProValMet-----SerAsnIleGlyLysSerThrAsnIleLeuTrp 148
 DB 424 GCGGAGAACAAAGTTTGGAAATGTCATCAATTGGCCGAGCGCTCCAAATCTCTGG 483
 QY 149 HisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuGlyGlnLysGlyCysVal 168
 DB 484 CATGATTGTCAGTTGGCAGCGCTGATCGCAGAGCTACTGAAGCAAGAGTTGCGTT 543
 QY 169 ValTrpIleThrGlyLeuSerGlySerGlySerThrLeuAlaCysAlaLeuSerArg 188
 DB 544 GTTGGATCACAGGACTTAGTGGTTTCAAGTAAAGTACCTGGCATGCACATTCAGATCG 603
 QY 189 GluLeuHisCysArgGlyHisLeuThrTrpValLeuAspGlyAspAsnLeuArgHisGly 208
 DB 604 GAGCTCCATACAAAGAGGGAAGCTTCTTATGTCTGTGATGGTGAATTTAAGACATGGT 663
 QY 209 LeuAsnArgAspLeuSerPhelysAlaGluAspArgAlaGluAsnIleArgArgValGly 228

DB 664 TTGAACAGGATCTTGGCTTTAAAGCGCGAAGACCGTCTGCTGCTGCTGCTGCTGCTGCT 723
 QY 229 GluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerPro 248
 DB 724 GAGGTAGCAAGAGCTATTTCGAGATGCAGCGCTAGTATGTCATTCGAAGTTTCATATCTCG 783
 QY 249 TyrArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSerAsnPheIleGluVal 268
 DB 784 TATAGGAGAGACCGTGGTCTTTCGTCATATTGTTCAGATGGTAGCTTTATTCAGTT 843
 QY 269 PheLeuAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeu 288
 DB 844 TTCTTGAACATCGCCCTTGGAAATGTGTGAGTCAAGGATCTCTAAGGGCTTTTATAAGCTT 903
 QY 289 AlaArgThrGlyLysIleLysGlyPheThrGlyLysAspAspProTyrGlyProPheIle 308
 DB 904 GCTCGTCAGGAAATAAAGGGTTTACTGGAATTTGATGACCCCTTATGAATCACCATT 963
 QY 309 AsnGlyGluIleValIleLysMetLysAspGluGluCysProSerProLysAlaMetAla 328
 DB 964 AACTCTGAGATTGAGATCAAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1023
 QY 329 LysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeu 341
 DB 1024 GGACAAGTAGTCACTTATCTAGAGGAGAAAGGGTTTCTG 1062

RESULT 3
 AF044285 1086 bp mRNA linear PUN 11-AUG-1998
 LOCUS Catharanthus roseus adenosine-5'-phosphosulfate-kinase (CRakn)
 DEFINITION mRNA, complete cds.
 ACCESSION AF044285
 VERSION AF044285.1 GI:2832299
 KEYWORDS Catharanthus roseus (Madagascar periwinkle)
 ORGANISM Catharanthus roseus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae;
 Vincet; Catharanthus.
 REFERENCE 1 (bases 1 to 1086)
 AUTHORS Arz, H.E., Gisselmann, G., Schiffmann, S. and Schwenn, J.D.
 TITLE A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana
 JOURNAL Biochim. Biophys. Acta 1218 (3), 447-452 (1994)
 MEDLINE 94325358
 PUBMED 8049272
 REFERENCE 2 (bases 1 to 1086)
 AUTHORS Schiffmann, S. and Schwenn, J.D.
 TITLE Isolation of cDNA clones encoding
 adenosine-5'-phosphosulfate-kinase (EC2.7.1.25) from Catharanthus
 roseus (Accession No. AF044285) and an isoform (akn2) from
 Arabidopsis (Accession No. AF043351) (PGR98-116)
 JOURNAL Plant Physiol. 117 (3), 1125 (1998)
 REFERENCE 3 (bases 1 to 1086)
 AUTHORS Schiffmann, S. and Schwenn, J.D.
 TITLE Direct Submission
 JOURNAL Submitted (23-JAN-1998) Biochemie der Pflanzen, Ruhr-Universität
 Bochum, Universitätsstrasse 150, Bochum 44780, Germany
 FEATURES
 source
 1. .1086
 /organism="Catharanthus roseus"
 /mol_type="mRNA"
 /db_xref="taxon:4058"
 /sub_clone="KSCRakn"
 /note="heterotrophic cell culture grown under sulfate
 deficiency"
 1. .1086
 /gene="CRakn"
 39. .977
 /gene="CRakn"
 /EC number="2.7.1.25"
 /function="phosphorylates 3'-OH group of
 adenosine-5'-phosphosulfate"

/note="APS-kinase; putative processing site after residue 46"
/codon_start=1
/product="adenosine-5'-phosphosulfate-kinase"
/protein_id="AAC31145.1"
/db_xref="GI:2832300"
/translation="MIGSVKRPVSVLCVPEFDFTSTGLGKSSSVKLPVNFAGFGSG
GVEKGLFAPTKATGSKTSFQVNGKVDNFRHLPQSDCNSSDLSLNNCGPFGKK
LQTTTVGNSNLIWHKCAVKSEROEPLQORGCVIWTGLSGSKSLTALCALSRGLH
AKGLTYILDGDNVHGLNSDLSPKADRAENIRIGEVAKLPADAGVICIASLSPY
RKPPDACKSLPEGFIVFMDVPLKVCARDPKGLYKLARAGKIKGTGIDDDPYEPP
LKSEIVLHQKMGMDSPCDLADIVISYLEENGYLKA"

ORIGIN

Alignment Scores:
Pred. No.: 1,6e-31 Length: 1086
Score: 833.00 Matches: 156
Percent Similarity: 84.11% Conservativeness: 24
Best Local Similarity: 72.90% Mismatches: 34
Query Match: 45.62% Indels: 0
DB: 8 Gaps: 0

US-10-829-432-4 (1-343) x AF044285 (1-1086)

Qy 130 ProVallysGluLysProValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHis 149
Db 333 CCAGGTAAAGATATTACAAACAACTACCGTTGGAACTCTACTAATATCTTTGGCAT 392

Qy 150 AsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuGlyGlnLysGlyCysValVal 169
Db 393 AATGTGCTGTGAAAGAAAGTGAAGGAGGAGCAACCCCTGCAGCAAGGGGCTGTGTATA 452

Qy 170 TrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGlu 189
Db 453 TGGATTACTGGTCTTAGTGGTTTCAGAAAGAGCACCTTGGCAATGCTTTAGTGCAGGC 512

Qy 190 LeuHisCysArgGlyHisLeuThrTrpValLeuAspGlyAspAsnLeuArgHisGlyLeu 209
Db 513 TTATCATGCAAGGAGAACTAACTACATCTTGTGATGATGATGATGATGATGATGATG 572

Qy 210 AsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgArgValGlyGlu 229
Db 573 AACAGTGTATCTTAGTGTATTAAGCAGAAAGATCGAGCGGAGATATTAGGAGGATTGAGAG 632

Qy 230 ValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuLeuSerProTyr 249
Db 633 GTTGCAAGCTCTTTGCTGATGCTGGAGTCATTTGATTCGAATGCAATGATTCCTCCCTAC 692

Qy 250 ArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSerAsnPheIleGluValPhe 269
Db 693 AGAAGCCCCCGATGCTGCTCTTGTCTTCCAGAGGGGATTTCAATGAGGTATTC 752

Qy 270 IleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAla 289
Db 753 ATGGATGCTCCCTCTGAAAGCTGCGAGGCAAGGACCCCTAAGGGCTTATATAAGCTTGA 812

Qy 290 ArgThrGlyLysIleLysGlyPheThrGlyIleAspAspProTyrGluProPheLeuAsn 309
Db 813 CGAGCTGGTAAATCAAGAGCTTTACTGGAATGATGATTCCTTACGAGCGCCACTTAAA 872

Qy 310 GlyGluIleValIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLys 329
Db 873 TCAGAGATTGTTCTACATCAGAAGCTTGGTATGTTGATGTTCTCCATGATGATTGGCTGAT 932

Qy 330 GlnValLeuCysTyrLeuGluLysGlyTyrLeuGlnAla 343
Db 933 ATAGTGATATCATATCTGGAGGAAATGGATATCTAAAGGCA 974

RESULT 4

LOCUS CQ805858 831 bp DNA linear PAT 10-MAY-2004
DEFINITION Sequence 2269 from Patent WO2004035798.
ACCESSION CQ805858

VERSION CQ805858.1 GI:47111563
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1
AUTHORS Inze, D., de Veylder, L. and Vlieghe, K.
TITLE Identification of novel e2f target genes and use thereof
JOURNAL Patent: WO 2004035798-A 2269 29-APR-2004;
CropDesign N.V. (BE)
FEATURES
source 1..831
location/Qualifiers
organism="Arabidopsis thaliana"
mol_type="unassigned DNA"
db_xref="taxon:3702"

ORIGIN

Alignment Scores:
Pred. No.: 7,94e-31 Length: 831
Score: 816.00 Matches: 159
Percent Similarity: 77.87% Conservativeness: 31
Best Local Similarity: 65.16% Mismatches: 44
Query Match: 44.69% Indels: 10
DB: 6 Gaps: 3

US-10-829-432-4 (1-343) x CQ805858 (1-831)

Qy 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsn 119
Db 127 GGATCTCAAACTCTGAGTCATAACAAAATGCTATCTCTGAGGTTAAATCCATTAAAC 186

Qy 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
Db 187 GGT-----CACCGGCAAAAGCAAGGACCT--TTGTCTACG 222

Qy 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db 223 GTCGGAACTCGACAAATATAAAGTGGCATGAATGTTCTGTGAGAAAGTTGATAGACAG 282

Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db 283 AGATTGCTTGATCAGAAAGGATGTGATTGGGTACCGGCTCTTAGTGGTTCAGGGAAG 342

Qy 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
Db 343 AGTACTTTGGCTTGTGCTTTTGAATCAGATGTTGATCAAAAGGGGAAGCTTTGTTATATT 402

Qy 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
Db 403 CTTGATGCTGATATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAAGCTGAGGAT 462

Qy 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 463 CGTCAGAGAATATTCTGATAGTGGAGGCTTGTCTAGAGCTTTTTCGCGATGCTGGAATA 522

Qy 240 IleCysIleAlaSerLeuLeuSerProTyrArgArgAspAspAlaCysArgAlaLeu 259
Db 523 ATCTGCATTCGAGTTTGTATCTCTCTATAGAACAGATAGGACGCTTGTGCAAGTTTG 582

Qy 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 593 CTCCCCGAGGAGAGATTTTGTGAGGTTCATGATGTACCGCTTAGTGTTCGAGGCG 642

Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db 643 AGGGATCCAAAGGGTCTTTTACAAGCTTCTCTGTCAGGAAAGATCAAAAGGTTTACCCTGG 702

Qy 300 IleAspAspProTyrGluProGluProGluIleValIleLysMetLysAspGlu 319
Db 703 ATCGATGACCTTACGAGCCACCATTTGAATGCGGAG-----ATTCTCTAGGACGTGAA 756

```

Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 757 GCAGGAACCTTCTCTATCGAAATGCGGAAAGGTCGTCGGATACCTTAGATAACAAGGCT 816

Qy 340 TyrLeuGlnAla 343
Db 817 TATCTTCAAGCA 828

RESULT 5
AX412633
LOCUS AX412633 831 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 397 from Patent WO0222675.
ACCESSION AX412633
VERSION AX412633.1 GI:21445091
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Glazebrook, J., Wang, X., Dangl, J. L., Eulgem, T. and Zhu, T.
TITLE Plant genes, the expression of which are altered by pathogen
infection
JOURNAL
Patent: WO 0222675-A 397 21-MAR-2002;
Syngenta Participations AG (CH); UNIVERSITY OF NORTH CAROLINA AT
CHAPEL HILL (US); Glazebrook, Jan (US); Wang, Xun (US); Dangl,
Jeffrey L. (US); Eulgem, Thomas (US)
AX412633
FEATURES
source
1..831
location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 7,94e-31 Length: 831
Score: 816.00 Matches: 159
Percent Similarity: 77.8% Conservative: 31
Best Local Similarity: 65.1% Mismatches: 44
Query Match: 44.6% Indels: 10
DB: 6 Gaps: 3

US-10-829-432-4 (1-343) x AX412633 (1-831)

Qy 100 GlyAlaArgThrHisCysHisArgGlyLeGlyArgTrpValArgArgArgAsn 119
Db 127 GGATCTCAAACTCTGAGTCATCAACAAAATGGATCTTCTCTGAGGTAAATCCATTAAAC 186

Qy 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
Db 187 GGT-----CACACGGGACAAAGCAAGGACCT---TTGTCTACG 222

Qy 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db 223 GTCGGAACCTCGACAAATATAAAGTGGCATGAATGTTCTGTGAGAAAGTTGATAGACAG 282

Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db 283 AGATTGCTTGATCAGAAAGGATGTGTGATTTGGGTCCCGGCTTTAGTGGTTACGGGAAG 342

Qy 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
Db 343 AGTACTTTGGCTGTGCTTGGATCAGATGTTGTATCAAAAGGGGAGGACCTTTGTTATATT 402

Qy 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
Db 403 CTTGATGGTGATAATGTTAGGCATGGCTTTAAACCGGTGATCTTTAGCTTTAAAGCTGAGGAT 462

Qy 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 463 CGTCGAGAGAATATTTCGTAGAGTTGGAGAGGTTGTAAAGCTTTTTCGCGAGTCTGGAATA 522

```

```

Qy 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
Db 523 ATCTGCATTGCGAGTTTGATATCTCTTATAGAACAGATAGGACGCTTGTGCAAGTTTG 582

Qy 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 583 CTCCCGCAGGAGGATTTTGTGAGGTGTTTCATGGATGTACCGCTTAGTGTTCGAGCGG 642

Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLeuLysGlyPheThrGly 299
Db 643 AGGATCCAAAGGCTCTTACAGCTTCTCGTCGAGGAAGATCAAGGTTTTCACCGGG 702

Qy 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMethLysAspGlu 319
Db 703 ATCGATGACCTTAGGAGCCACCATTTGAATGCGAG-----ATTCTCTAGGACGTGAA 756

Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 757 GGAGGAACCTTCTCTATCGAAATGCGGAAAGGTCGTCGGATACCTTAGATAACAAGGCT 816

Qy 340 TyrLeuGlnAla 343
Db 817 TATCTTCAAGCA 828

RESULT 6
AX506015
LOCUS AX506015 831 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 710 from Patent WO0216655.
ACCESSION AX506015
VERSION AX506015.1 GI:23387252
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Harper, J. F., Krebs, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 710 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)
FEATURES
source
1..831
location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores:
Pred. No.: 7,94e-31 Length: 831
Score: 816.00 Matches: 159
Percent Similarity: 77.8% Conservative: 31
Best Local Similarity: 65.1% Mismatches: 44
Query Match: 44.6% Indels: 10
DB: 6 Gaps: 3

US-10-829-432-4 (1-343) x AX506015 (1-831)

Qy 100 GlyAlaArgThrHisCysHisArgGlyLeGlyArgTrpValArgArgArgAsn 119
Db 127 GGATCTCAAACTCTGAGTCATCAACAAAATGGATCTTCTCTGAGGTAAATCCATTAAAC 186

Qy 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
Db 187 GGT-----CACACGGGACAAAGCAAGGACCT---TTGTCTACG 222

Qy 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db 223 GTCGGAACCTCGACAAATATAAAGTGGCATGAATGTTCTGTGAGAAAGTTGATAGACAG 282

Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179

```

```

Db      283 AGATTGCTTGATCAGAAAGGATGTGTGATTGGTTCACCGGTCTTAGTGGTTACAGGAAG 342
Qy      180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
Db      343 AGTACTTTGGCTGTGCTTGAATCAGATGTTGTATCAAAAGGGGAAGCTTTGTTATATT 402
Qy      200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLeuAlaGluAsp 219
Db      403 CTTGATGTTGATAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAGCTGAGGAT 462
Qy      220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db      463 CGTCGACAGAAATATTTCGTAGAGTTGGAGAGGTTGCTAAGCTTTTTCGCGATGCTGAATA 522
Qy      240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
Db      523 ATCTGCATTCGCGAGTTGATATCTCTCTATAGAACAGATAGGACGCTTGTGCAAGTTTG 582
Qy      260 LeuProHisSerAsnPhelIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db      583 CTCCTCCCGGAGAGATTTGTTGAGGTGTTTCATGGATGTACCGCTTAGTGTTCGAGCGG 642
Qy      280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db      643 AGGGATCCAAAGGCTCTTTACAAAGCTTGTCTGTCGAGGAAGATCAAAAGGTTTACCGGG 702
Qy      300 IleAspAspProTyrGluProProlIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db      703 ATCGATGACCCCTTACAGCCCATTAAGTCGAG-----ATTCTCTAGGACGTGAA 756
Qy      320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db      757 GGAGNACTTCTCCTATCGAATGCGGGAAGGTCGTCGGATATTAGATAACAAGGGT 816
Qy      340 TyrLeuGlnAla 343
Db      817 TATCTTCAAGCA 828

RESULT 7
AY132010
LOCUS      AY132010
DEFINITION Arabidopsis thaliana At2g14750/F26C24.11 mRNA, complete cds.
ACCESSION AY132010
VERSION    AY132010.1 GI:22135772
KEYWORDS   FLI CDNA.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
REFERENCE  1 (bases 1 to 831)
AUTHORS    Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE      Arabidopsis ORF clones
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 831)
AUTHORS    Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE      Direct Submission
JOURNAL    Submitted (26-JUN-2002) Salk Institute Genomic Analysis Laboratory

```

```

(SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PCEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Banh,J., Bowser,L., Chang,E., Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.
Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source             1..831
                        /organism="Arabidopsis thaliana"
                        /mol_type="mRNA"
                        /db_xref="taxon:3702"
                        /chromosome="2"
                        /clone="U13881"
                        /ecotype="Columbia"
                        /note="This clone is in pUNI 51"
     CDS                1..831
                        /note="putative adenosine phosphosulfate kinase"
                        /codon_start=1
                        /product="At2g14750/F26C24.11"
                        /protein_id="AA091043.1"
                        /db_xref="GI:22135773"
                        /translation="MIAAGAKSLGLSMAGPKGIFDSNSMSNSRVSVVVRACVSMDS
                        QTLSHNKGSIPEVKSINGHTQKQGPLSTVGNSTIKWHECSVEKVDRLQKGC
                        VIVTGLSGSGKSLTACALNMLYQKGLCYILDGNVRHGLNDRDLSPKADRAENIR
                        RVGEVAKLFADAGIICITASLIISPYTRDRDACRSLLPRGDFVEFMDVPLSVCEARDPK
                        GLYKLARAGKIKGFTGIDDDPEPLNCEISLGRGGTSPTEMAEKVGVGLNDKGYLQA
                        "
ORIGIN
Alignment Scores:
Pred. No.:          7.94e-31          Length:          831
Score:              816.00            Matches:         159
Percent Similarity: 77.87%            Conservative:    31
Best Local Similarity: 65.16%          Mismatches:     44
Query Match:        44.69%            Indels:         10
DB:                 8                  Gaps:           3
US-10-829-432-4 (1-343) x AY132010 (1-831)
Qy      100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsn 119
Db      127 GGATCTCAACTCTGAGTCATACAAAATGGATCTATTCTCGAGGTTAAATCCATTAAAC 186
Qy      120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
Db      187 GGT-----CACCGGGACAAAGCAAGGACCT---TTGTCTACG 222
Qy      140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db      223 GTCGGAACTCGACAAATATAAAGTGGCATGAATGTTCTGTGTGAGAAAGTTGATAGACAG 282
Qy      160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db      283 AGATTGCTTGATCAGAAAGGATGTTGTGATTTGGGTCACCGGCTCTTAGTGGTTACAGGAAG 342
Qy      180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
Db      343 AGTACTTTGGCTTGTGCTTGAATCAGATGTTGTATCAAAAGGGGAAGCTTTGTTATATT 402

```

sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Geneset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

Location/Qualifiers
 1..1068
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="14216"
 94..924
 /codon_start=1
 /product="putative adenosine phosphosulfate kinase"
 /protein_id="AA062496.1"
 /db_xref="GI:21553403"
 /translation="MIAAGAKSLGLSMASPKGIFDSNMSNSRSVVVVRACVSDGS QTLSHNKGSTPEVKSINGHTQKQKGLCYILDGNDVVRHGLNDRDLSPKADRAENIR VVGTGLSGSKGLACALNOMLYQKGLCYILDGNDVVRHGLNDRDLSPKADRAENIR RVGSAKLFPADAGIICIASLISPYRTDRCRSLIPSGDFVEFMDVPLSPCEARDPK GLYKLARAGKIKGFTGIDDDPYEPPLNCEISLREGGTSPIEMAEKVGVLDNKGYLQA"

ORIGIN

Alignment Scores:
 Pred. No.: 1,02e-30 Length: 1068
 Score: 816.00 Matches: 159
 Percent Similarity: 77.87% Conservative: 31
 Best Local Similarity: 65.16% Mismatches: 44
 Query Match: 44.69% Indels: 10
 DB: 8 Gaps: 3

US-10-829-432-4 (1-343) :x AY085264 (1-1068)

Qy 100 GlyAlaArgThrHisCysHisArgGlyLeuGlyArgTrpValArgArgArgAen 119
 Db 220 GGATCTCAAACTCTGAGTCATAACAAAATGATCTATCTCTGAGGTAAATCCATTAC 279
 Qy 120 GlyAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAen 139
 Db 280 GGT-----CACACGGGCAAAAGCAAGGACCT---TTGTCTACG 315
 Qy 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
 Db 316 GTCCGAACTCGACAAATATAAGTGGCATGAATGTTCTGTGTGAGAAAGTTGATAGACAG 375
 Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
 Db 376 AGATTGCTTGATCAGAAAGGATGTGATTTGGGTCCACCGCTCTAGTGGTTCAGGGAAG 435
 Qy 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
 Db 436 AGTACTTTGGCTTGTGCTTGAATCAGATGTTGTATCAAAAGGGAAGCTTTGTATATT 495
 Qy 200 LeuAspGlyAspAenLeuArgHisGlyLeuAsnArgHisLeuSerPheLeuAlaGluAsp 219
 Db 496 CTTGATGGTGATAATGTTAGGCATGGCTTAAACCGTGATCTTACTTTAAAGCTGAGGAT 555
 Qy 220 ArgAlaGluAenIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
 Db 556 CGTCAGAGATATTCGTAGAGTTGGAGAGTTGCTTAAGCTTTTTCGGATGCTGGAATA 615
 Qy 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
 Db 616 ATCTGCATTCGAGCTTTGATATCTCTTATAGAACAGATAGGACGCTTGTGCAAGTTTG 675
 Qy 260 LeuProHisSerAenPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
 Db 676 CTCCTCCCGAGGAGATTTTGTGTAGGTGTTTCATGATGTACCGCTTAGTGTTCGAGGCG 735
 Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299

Qy 200 LeuAspGlyAspAenLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
 Db 403 CTTGATGGTGATAATGTTAGGCATGCTTAAACCGTGTCTTAGCTTAAAGCTGAGAT 462
 Qy 220 ArgAlaGluAenIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
 Db 463 CGTCAGAGATATTCGTAGAGTTGGAGAGTTGCTTAAGCTTTTTCGGATGCTGGAATA 522
 Qy 240 IleCysIleAlaSerLeuIleSerProTyrArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 259
 Db 523 ATCTGCATTCGAGCTTTGATATCTCTTATAGAACAGATAGGACGCTTGTGCAAGTTTG 582
 Qy 260 LeuProHisSerAenPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
 Db 583 CTCCTCCCGAGGAGATTTTGTGTAGGTGTTTCATGATGTACCGCTTAGTGTTCGAGGCG 642
 Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
 Db 643 AGGATCCAAAGGCTTTTACAAAGCTTGTCTGTCAGGAAAGATCAAAAGGTTTACCGGG 702
 Qy 300 IleAspAspProTyrGluProIleLeuGlnGlyGluIleValIleLysMetLysAspGlu 319
 Db 703 ATCGATGACCTTTACGAGCCCATGTGAACCTCGAG-----ATTCTCTAGGACGTGAA 756
 Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
 Db 757 GGAGGACCTTCTCTATCAAAATGCGGAAAGGTCGTCGGATATCTAGATAACAAGGCT 816
 Qy 340 TyrLeuGlnAla 343
 Db 817 TATCTTCAAGCA 828

RESULT 8

AY085264 1068 bp mRNA linear PLN 14-APR-2003
 Arabidopsis thaliana clone 14216 mRNA, complete sequence.
 DEFINITION
 ACCESSION AY085264
 VERSION AY085264.1 GI:21403974
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1068)
 Haas,B.J., Volkovskiy,N., Town,C.D., Trukhan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.B.
 Full-length messenger RNA sequences greatly improve genome annotation
 Genome Biol. 3 (6), RESEARCH0029 (2002)
 22088475
 2 (bases 1 to 1068)
 Brover,V., Trukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
 Full-length cDNA from Arabidopsis thaliana
 Unpublished
 3 (bases 1 to 1068)
 Brover,V., Trukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
 Direct Submission
 Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA
 This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A

Db 736 AGGGATCCAAAGGGCTTTACAAAGCTTGTCTGTCAGGAAGATCAAAAGGTTTACCGGG 795
Qy 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db 796 ATCGATGACCCCTTACGAGCCACCATTTGAACTCGAG-----ATTCTCTAGGACGTGAA 849
Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 850 GGAGAACTTCTCCTATCGAATGCGGAAGAGTTCGTCGGATACTTAGATAACAAAGGT 909
Qy 340 TyrLeuGlnAla 343
Db 910 TATCTTCAAGCA 921

RESULT 9

ATU05238 1077 bp mRNA linear PLN 23-JUN-1998
LOCUS Arabidopsis thaliana APS kinase mRNA, complete cds.
DEFINITION U05238
ACCESSION U05238.1 GI:450234
VERSION
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1077)
Jain, A. and Leustek, T.
A cDNA clone for 5'-adenylylphosphosulfate kinase from Arabidopsis
thaliana

Plant Physiol. 105 (2), 771-772 (1994)
94345022
8066145
2 (bases 1 to 1077)
Lee, S. and Leustek, T.
APS kinase from Arabidopsis thaliana: genomic organization,
expression, and kinetic analysis of the recombinant enzyme
Biochem. Biophys. Res. Commun. 247 (1), 171-175 (1998)

98300303
9636674
3 (bases 1 to 1077)
Leustek, T.
Direct Submission
Submitted (18-JAN-1994) Thomas Leustek, Center for Agricultural
Molec. Biology, Rutgers University, Cook College, College Farm
Road, New Brunswick, NJ 08903, USA
Location/Qualifiers
1. .1077
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="ATK"
/tissue_type="whole plant"
33. .863
/EC_number="2.7.1.25"
/function="phosphorylates 5'-adenylylsulfate on
3'-hydroxyl group"

/note="similar to yeast APS kinase (MET14): SwissProt
Accession Number Q02196; similar to Escherichia coli APS
kinase (cysC): SwissProt Accession Number P23846;
adenosine 5'-phosphosulfate kinase"
/codon_start=1
/evidence=experimental
/product="APS kinase"
/protein_id="AAC50035.1"
/db_xref="GI:450235"

FEATURES

source

CDS

OTLSHKNKGSIPKVSINGHTGKOGPLSTVGNSTNIKWHECSVEKVDRLDQKGC
VTWVGLSGSGKSTLACALNOMLYOKGLCYILDGDNVHRHGLNRLDSFKADRAENIR
RVGEVAKLFADAGLICIASLISPYTRDRDACSLLPEGDFVEVMDVPLSCEARDPK
GLYKLARACKINGFTGIDDPYEPFLPNCISISLREGGTSPIEMAEKVGVLDNKGYLQA
"

transit_peptide 33. .143
/note="potential chloroplast transit peptide with a
protease cleavage site"
mat_peptide 144. .860
/product="APS kinase"
misc_feature 354. .377
/note="ATP/GTP binding site motif"

ORIGIN

Alignment Scores: 1.03e-30 Length: 1077
Pred. No.: 816.00 Matches: 159
Score: 77.87% Conservative: 31
Percent Similarity: 65.16% Mismatches: 44
Best Local Similarity: 44.69% Indels: 10
Query Match: 8 Gaps: 3
DB: 8
US-10-829-432-4 (1-343) x ATU05238 (1-1077)
Qy 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTyrValArgArgArgArgAsn 119
Db 159 GGATCTCAAACTCTGAGTCATAACAAAATGGATCTATTCTCTGAGGTTAAATCATCAATAC 218
Qy 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
Db 219 GGT-----CACACGGGACAAAAGCAAGGACCT---TTCTCTACG 254
Qy 140 IleGlyLysSerThrAsnIleLeuTyrHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db 255 GTCGGAACCTCGACAAATATAAAGTGGCATGAATGTTCTGTGAGAAAGTTGATAGACAG 314
Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTyrPileThrGlyLeuSerGlySerGlyLys 179
Db 315 AGATTGCTGATCAGAAAGGATGTGTGATTGGGTACCGGCTCTAGTGGTTCAGGGAAG 374
Qy 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
Db 375 AGTACTTTGGCTTGTGCTTTGAATCAGATGTTGTATCAAAAAGGGGAAAGCTTTGTATATT 434
Qy 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
Db 435 CTTGATGGTGATAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAAGCTTGAGGAT 494
Qy 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 495 CGTCACAGAAATATTCTGTAGAGTTGGAGAGGTTCTTAAGCTTTTTCGGGATGCTGGAATA 554
Qy 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
Db 555 ATCTGCATTGCGAGTTTGTATCTCTTATAGAACAGATAGGACGCTTGTGCAAGTTTG 614
Qy 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 615 CTCCCCGAGGAGAGATTTTGTTGAGGTGTTCAATGATGTACCGCTTAGTGTGTTGAGGCG 674
Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db 675 AGGATCCAAAGGCTTTTACAAGCTTGCTGTCAGGAAAGATCAAAAGGTTTACCGGG 734
Qy 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db 735 ATCGATGACCCCTTACGAGCCACCATTTGAACCTGCAG-----ATTCTCTAGGACGTGAA 788
Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 789 GGAGAACTTCTCCTATCGAATGCGGAAGAGTTCGTCGGATACTTAGATAACAAAGGT 848
Qy 340 TyrLeuGlnAla 343
Db 849 TATCTTCAAGCA 860
RESULT 10
AT054287

```

LOCUS      AY054287      1143 bp      mRNA      linear      PLN 30-SEP-2001
DEFINITION Arabidopsis thaliana At2g14750/F26C24.11 mRNA, complete cds.
ACCESSION  AY054287
VERSION     AY054287.1  GI:15810037
KEYWORDS    FLI CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1143)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Arabidopsis cDNA clones
2 (bases 1 to 1143)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Unpublished
TITLE      Direct Submission
JOURNAL
REFERENCE  Submitted (22-AUG-2001) Salk Institute Genomic Analysis Laboratory
          (SIGNAL), plant Biology Laboratory, The Salk Institute for
          Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
          USA
COMMENT    RIKEN Genomic Sciences Center (GSC) members carried out the
          collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
          Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
          Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
          Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bower,L.,
Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.
Location/Qualifiers
1..1143
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosomes="2"
/clone="RAFL09-19-L13 (R13881)"
/ecotype="Columbia"
1..92
93..923
/note="putative adenosine phosphosulfate kinase"
/codon_start=1
/product="At2g14750/F26C24.11"
/protein_id="AA06946.1"
/db_xref="GI:15810038"
/translacion="MTAGAKSLGLSWASPKGIFDSSMSNRSVVVVRACVMDGSG
QTLHNKSGSIFEVKSGHTGCKQKQPLSTVGNSTNKNHSCSVEKVRQRLQKGC
V1WVTGLSGSGKSLACLANQMLYOKGLCYLLDGNVNRHGLNLDLSFKABDRANR
RYGEVAKFADAGLIICTASLSPYTRDRCASLTPEDGFVEFMDVPLSCVCEARDPK
GLYLKARAKIGFTGIDDPVEPPLNCBISLREGGTSPIEMAEKVVVGLDNKGYLQA

FEATURES             source
source               source
5'UTR
CDS

```

```

3'UTR
ORIGIN
Alignment Scores:
Pred. No.:      1.09e-30      Length:      1143
Score:          816.00      Matches:    159
Percent Similarity: 77.87%      Conservative: 31
Best Local Similarity: 65.16%      Mismatches: 44
Query Match:    44.69%      Indels:    10
DB:            8      Gaps:    3
US-10-829-432-4 (1-343) x AY054287 (1-1143)
QY 100 GYAlaAAGThrHisCysHisArgGlyLeuGlyArgTdpValArgArgArgArgaen 119
Db 219 GGAUCTCAAACTCGAGTCATAACAAAATGATCTATTCTCGAGGTAAATCCATTAAAC 278
QY 120 GYAlaAAlaProGlyGluAlaProHisSerProValHisGluSerProValMetSerAen 139
Db 279 GGT-----CACACGGGACAAAAGGAGGACCT--TTGCTACG 314
QY 140 lLeGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db 315 GTCCGAAACTCGACAAATATAAAGTGGCATGAATGTTCTGTGAGAAAAGTTGATAGACAG 374
QY 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db 375 AGATTGCTTGATCAGAAAGGATGTGATTGGGTGACCGGTCTTAGTGGTTCAGGGAAG 434
QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyVal 199
Db 435 AGTACTTTGGCTTGTGCTTTGATCAATCAGATGTTGATCAAAAGGGGAGGCTTTGTTATAT 494
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAep 219
Db 495 CTTGATGCTGATAAATTTAGGCGATGGCTTAAACCGTGATCTTAGCTTTAAAGCTGAGGAT 554
QY 220 ArgAlaGluAsnIleAatGatGValGlyValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 555 CGTCAGAGAAATATTCGTAGAGTTGGAGAGGTTCGTAAGCTTTTTCGGGATGCTGGAATA 614
QY 240 lLeCysIleAlaSerLeuIleSerProTyArgArgAspArgAspAlaCysArgAlaLeu 259
Db 615 ATCTGCATTGCGAGTTGATATCTCCATTATAGACAGATAGGACGCTTGTCGAGTTTG 674
QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 675 CTCCCGCAGGAGGAGATTTTGTGAGGTTCATGATGTACCGCTTAGTGTTCGAGGCG 734
QY 280 ArgAspProLysGlyLeuTyLysLeuAlaAargThrGlyLysIleLysGlyPheThrGly 299
Db 735 AGGATCCAAAGGGTCTTTTACAAGCTGCTCGTCAGGAAAGATCAAAAGGTTTATCCCGGG 794
QY 300 lLeAspAspProTyGluProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db 795 ATCGATGACCTTACGAGCCACCATGTAATGCCAG-----ATTCTCTAGACGCTGAA 848
QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyLysLeuGluGluAenGly 339
Db 849 GGAGGAATCTTCTCTATCGAAATGCGGAAAGGTCGTGCGGATACCTTAGATAACAAGGCT 908
QY 340 TyrLeuGlnAla 343
Db 909 TATCTTCAAGCA 920
RESULT 11
LOCUS      ATCLAPSK      1185 bp      mRNA      linear      PLN 05-FEB-1998
DEFINITION A. thaliana (L. Heynh.) chloroplast mRNA for recombinant APS-kinase.
ACCESSION  X75782
VERSION     X75782.1  GI:414736
KEYWORDS    APS-kinase; ATP:adenylylsulfate-3'-phosphotransferase.

```

SOURCE ORGANISM	REFERENCE
Arabidopsis thaliana (thale cress)	1
Arabidopsis thaliana	Arz, H.E., Gisselmann, G., Schiffmann, S. and Schwenn, J.D.
Eukaryota; Viridiplantae	A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana
Eukaryota; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	Biochim. Biophys. Acta 1218 (3), 447-452 (1994)
	94325358
	PUBMED
	8049272
	(sites)
	2 (bases 1 to 1185)
	Arz, H.E., Gisselmann, G., Schiffmann, S. and Schwenn, J.D.
	A chloroplast APS-kinase cDNA from Arabidopsis thaliana
	Unpublished
	3 (bases 1 to 1185)
	Schwenn, J.D.
	Direct Submission
	Submitted (03-NOV-1993) Schwenn J. D., Ruhr University Bochum, Biology, Universitaetstr. 150, 44780 Bochum, Germany
	Location/Qualifiers
	FEATURES

FEATURES
SOURCE

```

source
1. .1185
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="PM239x14cDNA"
/tissue_type="leaf"
/clone_lib="cDNA in lambda ZapII"
/dev_stage="vegetative"
42. .872
/EC_number="2.7.1.25"
/function="phosphorylates 3'-OH group of adenylylsulfate"
/standard_name="ATP:adenylylsulfate-3'-phosphotransferase"
/citation=[2]
/codon_start=1
/product="APS kinase"
/protein_id="CAA53426.1"
/db_xref="GI:414737"
/db_xref="GOA:Q43295"
/db_xref="UniProt/Swiss-Prot:Q43295"
/translation="MIAGAKSLGLSWASPKGIFDSNSMSNSRSVVVVVACVSMDSGS
QTLSHNKGSTPEVKSNGHTGQKQGLSPVTGNSNTNWKHECSVEKVDRORLDDQKGC
VIWVTGSGSKSTLACNLQMLYQKQKLCYILDGDNVRHGLRNDJLSFKAEADRNIIR
RVGEVAKLFDADAGICTASLSPYRTDRCASRLIPGDEGFVMDPLVPLSCVSPARDPK
GLYLAARAGKIKGFTPTDGLPPLNCEIISLGREGGTSPIEMAEKVVGVDLNDKGYLQA

```

mat peptide

```
mat_peptide
/product="APS-kinase"
/EC number="2.7.1.25"
/notes="putative mature peptide only - gene product
starting with ATG at position 162 is enzymatically active,
but maturation in the chloroplast has not yet been shown.
Homology to sequences from E. coli & ye"
/citation={1}
/citation={2}
```

```

/ CIRCULATION=[2]
transit peptide 42: .153

```

```

stem loop      984. .1012
polyA_site     985. .990
stem loop     1092. .1120
/notes="putative transit peptide with a processing site
suitable for stromal protease; ref. Gavel & Heijne FEBS
Lett. 261, 455, 1990"

```

ORIGIN

Alignment Scores:	1,13e-30	Length:	1185
Pred. No.:	816.00	Matches:	159
Score:	77.8%	Conservative:	31
Percent Similarity:	65.16%	Mismatches:	44
Best Local Similarity:	44.65%	Indels:	10
Query Match:	8	Gaps:	3
DB:			

US-10-829-432-4 (1-343) x ATCLAPSK (1-1185)

Qy	100	GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsn	119
Db	168	GGATCTCAAACTCTGAGTCATAACAAAATGGATCTATTCTCTGAGGTAAATCCATTAAAC	227
Qy	120	GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn	139
Db	228	GGT-----CACCGGGAACAAAGCAAGGACCT---TTGTCTTACG	263
Qy	140	IleGlyLysSerThrAsnIleLeuTyrHisAsnCysLeuIleGlyGlnSerAspArgGln	159
Db	264	GTGGAATACTGCACAAATATAAGTGGCATGAATGTTCTTGAGAAAGTTGATAGACAG	323
Qy	160	LysLeuLeuGlyGlnLysGlyCysValValTyrIleThrGlyLeuSerGlySerGlyLys	179
Db	324	AGATTGCTTGATCAGAAAGGATGTGTATGGGTACCGCTCTTAGTGTTTCAGGGAAG	383
Qy	180	SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal	199
Db	384	AGTACTTTGCTTGTGCTTTGAATCAGATGTTGTATCAAAAGGGGAAGCTTTGTTATATT	443
Qy	200	LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp	219
Db	444	CTTGATGGTGATAATGTAGCATGGCTTAAACCGTGATCTTAGCTTTAAAGCTGAGGAT	503
Qy	220	ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAlaGlyVal	239
Db	504	CGTGCAGAGAATATTCGTAGAGTTGGAGAGGTTGCTAAGCTTTTTCGGATGCTGGAATA	563
Qy	240	IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu	259
Db	564	ATTCGATTCGGAGTTTGATATCTCTTATAGAACAGATAGGACGCTTGTCGAAGTTTG	623
Qy	260	LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla	279
Db	624	CTCCCCGAGCGAGATTTTGTGTAGGTCCTCATGGATGTCACCGCTTAGTGTTCGCGAGCG	683
Qy	280	ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly	299
Db	684	AGGATCCAAAGGTCCTTACAAGCTTCGTCTGTCGAGAAAGATCAAGGTTTTCACGGG	743
Qy	300	IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu	319
Db	744	ATCGATGACCTTTACGAGCCACCATTTAACTGCAG-----ATTTCTTAGGACGCTGAA	797
Qy	320	GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrIleuGluGluAsnGly	339
Db	798	GGAGGAACCTTCTCTATCGAAATGGCGGAAAGGTCGTGGATACTTAGATACAAAGGCT	857
Qy	340	TyrLeuGlnAla	343
Db	858	TATCTTCAACA	869
RESULT 12	AF178976	970 bp	linear
LOCUS	AF178976	Zea mays adenosine-5'-phosphosulfate kinase (AK1)	PLN 13-DEC-1999
DEFINITION	Zea mays adenosine-5'-phosphosulfate kinase (AK1) mRNA, partial cds.		
ACCESSION	AF178976		
VERSION	AF178976.1	GI:6563284	
KEYWORDS			
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 970)		
TITLE	Bolchi,A., Petruccio,S. and Ottomello,S.		
JOURNAL	Isolation and comparative expression analysis of a maize cDNA encoding adenosine 5'-phosphosulfate kinase		
REFERENCE	Unpublished		
	2 (bases 1 to 970)		

AUTHORS Bolchi,A., Petrucco,S. and Ottonello,S.
 TITLE Direct Submission
 JOURNAL Submitted (19-AUG-1999) Istituto di Scienze Biochimiche,
 Universita' di Parma, Viale delle Scienze, Parma 43100, Italy

FEATURES
 source
 1..970
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="Dekalb XL72"
 /db_xref="taxon:4577"
 <1..970
 /gene="AK1"
 <1..868
 /gene="AK1"
 /EC number="2.7.1.25"
 /note="APS kinase; ATP adenosine-5'-phosphosulfate
 3'-phosphotransferase"
 /codon_start=2
 /product="adenosine-5'-phosphosulfate kinase"
 /protein_id="AAF17236.1"
 /db_xref="GI:6563285"
 /translation="VVPVLLLAVARASRAPEPEPRVKRSVASAPPPPPETPPGQWQ
 NGDRRRLKRLGPELVECTGDRSVEQPEHAGVEGKAWKMSSTVPKSSNIFWHDPCVKG
 TRQNVLKQKGVVITGUSGSKSTLACTLREHLRGKLAIVLDGNLRHGLNKL
 GPKAEDRNIIRVAKLPADAGLVCIATLSIHPHRDRESCLALSDSFIEVFLNMS
 LELCEARDPKGLYKLARAGIKGFTGIDDPYEAPLNCEIEIKVEDVGVCPPPAEMAGQV
 VTYLBEKGFLHE"

ORIGIN
 Alignment Scores:
 Pred. No.: 9,78e-31 Length: 970
 Score: 815.50 Matches: 178
 Percent Similarity: 67.97% Conservative: 30
 Best Local Similarity: 58.17% Mismatches: 65
 Query Match: 44.66% Indels: 33
 DB: 8 Gaps: 6

US-10-829-432-4 (1-343) x AFI78976 (1-970)

QY 40 ProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProProValLeuPro 59
 DB 33 CCGCGCCCTCGCGTCGCGCACCTG-----AACCGAGGCCCCCGCTCAAAATCTA 80
 QY 60 GlyLeuThrProSerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrProArg 79
 DB 81 GGGTTGCATCGCGCGCCCCCTCTCCGAGACTCCACGCGCTGGAACCAAGAGGAGG 140
 QY 80 SerSerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgGluGlyGluGlyArg 99
 DB 141 GTG-----ACAGCGCAGCAAGAGCTGCGAGGAGCGC 170
 QY 100 GlyAla-ArgThrHisCysHisArgGlyLeuGlyArgTrpValArgArgArgArgAs 119
 DB 171 CCTCGTCGAATGCAC-----TGGCGACCGCTCCGTCGA----- 204
 QY 119 nGlyAlaAlaProGlyGluAla-ProHisSerProValLysGluLys-----ProValM 137
 DB 205 -----GGAGCAGCGGAGCAGCAGGAGTTCAGGAGAAAGCTTGGAAATGT 251
 QY 137 etSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerA 157
 DB 252 CATCCACCGTCGCGAAGTCATCAAAATATCTTCGCGACGATTGTCAGTTGGCAAGACTG 311
 QY 157 spArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlyS 177
 DB 312 ATCGCCAGAACGACTCAAGACAGAAAGGCTGTGTGTCGTGATTACAGGCCCTTAGTGGTT 371
 QY 177 exGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuT 197
 DB 372 CAGGTAAAGTAGTACCTTGGCATGTACATTAGCGCGTGAGCTCCACACAGAGGGAAGCTTG 431
 QY 197 hrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysA 217

Db 432 CATATGTCCTTGACCGTGTATTAACATTTAGACATGGTCTGAACAAGGATCTTGGCTTCAAG 491
 QY 217 laGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspA 237
 Db 492 CTGAAGACCGTGTCTGAAAATATAAGAGAGTT-----GTACCAAGTTATTATTGAGATG 545
 QY 237 laGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysA 257
 Db 546 CGGGCCCTTGTATGTTCAAGTTTATATCTCCACATAGGAGACCGTGAATCTTGGC 605
 QY 257 rgAlaLeuLeuProHisSerAsnPhelIleGluValPheIleAspLeuProLeuLysIleC 277
 Db 606 GTGCATTGTTGTCATAGTAGCTTCATTGAAGTTTCTTCAACATGCTCTTGGAAATTAT 665
 QY 277 ysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyP 297
 Db 666 GTGAAGCAAGAGATCCAAAGGCCCTCTATAGCTTGTCTGTCGACGAGAAAATAAAGGTT 725
 QY 297 heThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetL 317
 Db 726 TTACAGGAATAGATGACCTTTATGAAGCACCACTGAATTGTGAGATTGAGATCAAAGAGG 785
 QY 317 ysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluG 337
 Db 786 TAGATGGTGTATGCCCTCGCCCTGCTGAGATGGCAGGCAAGTTGTTACTTACCTTGAGG 845
 QY 337 luAsnGlyTyrLeu 341
 Db 846 AGAAGGCTTCTCTG 859

RESULT 13
 LOCUS BD248389 920 bp DNA linear PAT 17-JUL-2003
 DEFINITION Gene participating in tolerance against environmental stress.
 ACCESSION BD248389
 VERSION BD248389.1 GI:33058159
 KEYWORDS JP 2002524052-A/33.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
 REFERENCE 1 (bases 1 to 920)
 AUTHORS Lee, J.H. and Verbruggen, N.
 TITLE Gene participating in tolerance against environmental stress
 JOURNAL Patent: JP 2002524052-A 33 06-AUG-2002;
 CROPDESIGN NV
 COMMENT OS Arabidopsis thaliana (thale cress)
 PN JP 2002524052-A/33
 PD 06-AUG-2002
 PF 04-AUG-1999 JP 2000563810
 PR 04-AUG-1998 EP 98202634.6
 PI JHONG HEE LEE, NATHALIE VERBRUGEN
 PC C12N15/09, A01H5/00, C07K14/415, C12N5/10, C12N9/12, C12N15/00, PC
 C12N5/00
 CC Gene participating in tolerance against environmental stress
 FH Key Location/Qualifiers
 FT CDS
 Location/Qualifiers
 1..920
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"

ORIGIN
 Alignment Scores:
 Pred. No.: 5.39e-29 Length: 920
 Score: 778.50 Matches: 175
 Percent Similarity: 66.04% Conservative: 37
 Best Local Similarity: 54.52% Mismatches: 67
 Query Match: 42.63% Indels: 43
 DB: 6 Gaps: 9

```
US-10-829-432-4 (1-343) x BD248389 (1-920)
QY 40 ProThrProThrLeuAlaValIleLeuValAenProGlnArgAlaProProVal----- 57
Db 14 CCGTCGGATTCAAATGAAGATTAGCTATCAGAGCATCTCGA---CCGTGCGTTTTCGTGT 70
QY 58 ---LeuProGlyLeu-----ThrProSerAspAla-----Pro 67
Db 71 TCTATTCAGGTCCTCGCGCGATTCGCCAGAAACCTCCAAAGTGACGGTTCCTCAAG 130
QY 68 LeuProAlaLeuValIleHisGlyLeuThrProArgSerSerHisSerAlaGlyLeu 87
Db 131 CTGCTCGGTC-GTCTATTCCGCGGACAGCCGAAATTAGTAGCGAATCTTACTTCCT 189
QY 88 AlaSerAspSerGlyArgArgGluGlyGluGlyArgGlyAlaArgThrHisCysHisArg 107
Db 190 TCAATCAATCTCAGCCGT-----TAACTGTCTCTGCTCAAGTCTCCTCACCGC 237
QY 108 GlyIleGlyArg-----TrpValArgArg-ArgArgArgAsnGlyValAlaApr 123
Db 238 TGATTTTCCCGCCTTTCAGAACTATCTAGAAAGGGGAGAAATAACGGA----- 289
QY 123 oGlyGluAlaProHisSerProValLysGluLysProValMetSerAsnIleGlyLysSe 143
Db 290 -----AAAGAG----- 300
QY 143 rThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuGlu 163
Db 301 AGAGACATCTGTGTGCGACGAGTTCGATATGCGAGTCGACACACAACTCTTCA 360
QY 163 yGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAl 183
Db 361 ACAAAAGGGTCTGTCTGTGATCTACTGTCTCAGTGTCTCAGGAAAGACACTGTTC 420
QY 183 aCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrrValLeuAspGlyAs 203
Db 421 TTGTGCACTAAGTAAAGCAATCTTTTGAAGAGGCAAACTTACTTACACACTCGAGCGGA 480
QY 203 pAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAs 223
Db 481 CAATGTCGTCACGGCTTAAACCGTGACCTCTTCAAGCAGAGCACCAGCCGAAAA 540
QY 223 nIleArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAl 243
Db 541 CATTAGAAGAAATTTGTGAGGTGGCTAAGTTGTTGTCGACGTCGAGTCAATTTGTATAGC 600
QY 243 aSerLeuIleSerProTyrrArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSe 263
Db 601 AAGTTTGATTTCTCCGTACCGGAGACAGACGCGTCCCGTCTCTTTTACCTGACGG 660
QY 263 rAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProly 283
Db 661 CGATTTGTCGAGGCTTTCATGAGAGTCTCTCTTTCATGTGCGAGTCGAGATCCAAA 720
QY 283 sGlyLeuTyrrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspPr 303
Db 721 GGGGTTGTACAAGCTTGCACGTGCGAGGCAAAATCAAAAGGCTTCACTGGAATCGACGCC 780
QY 303 oTyrrGluProProIleAsnGlyGluIleValIleLys---MetLysAspGluGluCysPr 322
Db 781 TTACGAGGCACCATGTAATTCGAGAGTAGTGTCTGAAACACACAGGAGACGAGTCGTG 840
QY 322 oSerProLysAlaMetAlaLysGlnValLeuLysCysTyrrLeuGluGluAsnGlyTyrrLeuG 342
Db 841 TTCGCCACGTCAGATGGCTGAGAACATCATCTCTTACCTGCAAAACAAAGGTATCTTGA 900
QY 342 n 342
Db 901 G 901
RESULT 14
AF462823
LOCUS
```

```
DEFINITION Arabidopsis thaliana AT4g39940/T5J17_110 mRNA, complete cds.
ACCESSION AF462823
VERSION AF462823.1 GI:18087562
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1243)
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE Arabidopsis cDNA clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1243)
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-2001) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura, Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.
FEATURES
source
1..1243
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clones="RAFL09-09-101(R21470)"
/ecotype="Columbia"
70..951
/note="adenosine-5-phosphosulfate-kinase"
/product="AT4g39940/T5J17_110"
/protein_id="RAL58913.1"
/db_xref="GI:18087563"
/translation="MEGLAINRASRPSVFCSTPGLGGSHRKPDDGFLKLPASSIPAD SRKLVA NSTSFHPISAVNVAQASLTADFPALSETILKEGNNKBAENIVHESHSI CRCDROQLQKGCWVITGLSGSKSTVACALKALFERKGLTYTIDGNDVHGLNLR DLTFAEHTNIRRIGEIVAKLFADVGIVICIASLISPYRRDKRLDLPDGFVEFV MDVPLHWCESDPKGLYKLARAGIKGTGTGDDPYEAPVNCVVLKHTGDDDESCPRQ MAENIIISLQNKYLEG"
5'UTR
CDS
```

```

3' UTR      952..1243
ORIGIN
Alignment Scores:
Pred. No.:      7,27e-29      Length:      1243
Score:          778.50      Matches:      175
Percent Similarity: 66.04%      Conservative: 37
Best Local Similarity: 54.52%      Mismatches: 67
Query Match:      42.63%      Indels:      43
DB:              8          Gaps:      9

US-10-829-432-4 (1-343) x AF462823 (1-1243)
QY      40  ProThrProThrLeuAlaValleLeuValAsnProGlnArgAlaProVal----- 57
Db      58  CCGTCGGATTCATGAAGATTAGCTATCAGAGCATCTCGA---CCGTCGGTTTCTGT 114
QY      58  ----LeuProGlyLeu-----ThrProSerAspAla-----Pro 67
Db      115 TCTATTCCAGGTCCTCGCGCGGATTCACCGAAACCTCCAAAGTGACGGTTTCCTCAAG 174
QY      68  LeuProAlaLeuValleHisGlyLeuThrProArgSerSerHisSerAlaGlyLeu 87
Db      175  GTGCTCGGCTC-GTCTATTCCGCGCGGACAGCCGAAATAGTAGCGAATTCTACTTCCTT 233
QY      88  AlaSerAspSerGlyArgArgGlyGlyGlyArgGlyAlaArgThrHisCysHisArg 107
Db      234  TCATCCAAATCTACGCGT-----TAACTGCTCTGCTCAAGCTTCCTCACCOC 281
QY      108  GlyileGlyArg-----TrpValArgArg-ArgArgArgAsnGlyAlaAlaPr 123
Db      282  TGATTTTCCCGCCCTTTCAGAACTATATACGAAGAGGGAAGAAATAACGGA----- 333
QY      123  oclyGluAlaProHisSerProVallyGlyLysProValMetSerAsnileGlyLysSe 143
Db      334  ----AAAGAG-----AAAGC 344
QY      143  rThrAsnileLeuTrpHisAsnCysLeuileGlyGlnSerAspArgGlnLysLeuLeuG1 163
Db      345  AGAGAACATCGTGCGCAGAGATTCGATATGCAGATCGCAGACACACACTTCTTCA 404
QY      163  yGlnLysGlyCysValValPileThrGlyLeuSerGlySerGlyLysSerThrLeuAl 183
Db      405  ACAAAAGGGTGTGTGCTTGGATCTCAGTCTCAGTGTGTTTCTAGGAAAGCACTGTTGC 464
QY      183  aCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGlyAs 203
Db      465  TTGTGCACCTAAGTAAAGCAATTGTTTAAAGAGGCGCAAACTTACTTACACACTCGACGCGCA 524
QY      203  pAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAs 223
Db      525  CAATGTCCGTACGGGCTTAAACGTCACCTACTTTCAAAGCAGAGCACCAGCAGCAAAA 584
QY      223  nleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValileCysileAl 243
Db      585  CATTAGAACAATTGGTGAGTGGCTAAGTGTGTTGCTGACGCGGAGTCAATTGTATAGC 644
QY      243  aSerLeulleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSe 263
Db      645  AAGTTTGATTTCTCCGTACCGCAGAGACAGACAGACGCGTCCCGGCTCTGTTTACCTGACGG 704
QY      263  rAsnPheilleGluValPheileAspLeuProLeuLysilleCysGluAlaArgAspProly 283
Db      705  CGATTTCTGCGAGAGTCTTCATGAGCGTCTCTTCATGTGTCGAGTCCAGAGATCCAAA 764
QY      283  sGlyLeuTyrLysLeuAlaArgThrGlyLysilleLysGlyPheThrGlyLysileAspPr 303
Db      765  GGGTTGTACAAGCTTGCAGGTGCAGGCAAAATCAAAAGGCTTCCTGGAATCAGACGCC 824
QY      303  oTyrGluProProIleAsnGlyGluilleValleLys---MetLysAspGluGluCysPr 322
Db      825  TTACGAGGCACCAAGTGAATTGCGAGGTAGTGTGTAACACACACAGGAGACGACGAGTCGTG 884

```

322 oSerProLysAlaMetAlaLysGlnValleuCysTyrLeuGluGluAsnGlyTyrLeuG1 342

885 TTCGCCAGCTCAGATGGCTGAGACATCATCTCTTACTTGCACAAACAAAGGTTATCTTGA 944

342 n 342

945 G 945

RESULT 15

AF043351

LOCUS

DEFINITION

Arabidopsis thaliana adenosine-5'-phosphosulfate-kinase (akn2)

VERSION

AF043351.1

KEYWORDS

Arabidopsis thaliana (chale cress)

ORGANISM

Arabidopsis thaliana

REFERENCE

1 (bases 1 to 1311)

Arz,H.E., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.

A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana

Biochim. Biophys. Acta 1218 (3), 447-452 (1994)

MEDLINE

94325358

PUBMED

8049272

REFERENCE

2 (bases 1 to 1311)

Schiffmann,S. and Schwenn,J.D.

Isolation of cDNA clones encoding

adenosine-5'-phosphosulfate-kinase (BC2.7.1.25) from Catharanthus

roseus (Accession No. AF044285) and an isoform (akn2) from

Arabidopsis (Accession No. AF043351) (PCR98-116)

JOURNAL

Plant Physiol. 117 (3), 1125 (1998)

REFERENCE

3 (bases 1 to 1311)

Schiffmann,S. and Schwenn,J.D.

Direct Submission

Submitted (16-JAN-1998) Biochemie der Pflanzen, Ruhr-Universitaet

Bochum, Universitaetsstrasse 150, Bochum 44780, Germany

FEATURES

Location/Qualifiers

1..1311

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/cultivar="Columbia"

/db_xref="taxon:3702"

/clone="pGAD10"

/tissue_type="green vegetative"

/dev_stage="3 week old"

1..1311

/gene="akn2"

61..942

/gene="akn2"

/EC_number="2.7.1.25"

/function="phosphorylates 3'-OH group of

adenosine-5'-phosphosulfate"

/note="APS-kinase isoform; similar to the product encoded

by Genbank Accession Number X75782"

/codon_start=1

/product="adenosine-5'-phosphosulfate-kinase"

/protein_id="AAC39520.1"

/db_xref="GI:2829133"

/translation="MEGLAIRASRPSVFCSTPGLGDSHRKPPSPDGFLKLPASSIPAD

SRKLVANSTSPHPISAVNVSAQSLTADTPALSETILKEGRNNGKEKENVHSSI

CRCKRQLQKQCGCVMTITLUGSGSKSTVACALSKALPFRKLTITLQDGNVRHGLNR

DLTTPKAERTENIRRIEVAKLFDVGVICISLSPYRDRDRACRSLLPDGFVEVF

MDVPLHYCESRDPKGLYKLARAGIKGTGTDDPFEAPVNCVVLKHTGGDDESCSPRQ

MAENIISYLNKGYLEG"

238..240

/misc_feature

/gene="akn2"

/note="putative processing site 3' of this position;

unclassified site"

ORIGIN

Alignment Scores:

Pred. No.: 7,67e-29 Length: 1311
Score: 778.50 Matches: 175
Percent Similarity: 66.04% Conservative: 37
Best Local Similarity: 54.52% Mismatches: 67
Query Match: 42.63% Indels: 43
DB: 8 Gaps: 9

US-10-829-432-4 (1-343) x AF043351 (1-1311)

```
QY 40 ProThrProLeuAlaValIleHisGlyLeuThrProArgSerSerHisSerSerAlaGlyLeu 57
||||| : : : : :
Db 49 CCGTCGGATTCAATGAGGATTAGCTATCAGAGCATCTCGA---CCGTCGGTTTCTGT 105
QY 58 ---LeuProGlyLeu-----ThrProSerAspAla-----Pro 67
||||| : : : : :
Db 106 TCTATTCAGGTCCTCGCGCGGATTCACCCAGAAACCTCCAAAGTGACGGTTTCTCAAG 165
QY 68 LeuProAlaLeuValIleHisGlyLeuThrProArgSerSerHisSerSerAlaGlyLeu 87
||||| : : : : :
Db 166 CTGCTGGGTC-GTCTATTCCGGCGGACAGCCGAAATTAGTAGCGAATTCTACTTCCTT 224
QY 88 AlaSerAspSerGlyArgGluGlyGluGlyAlaArgGlyAlaArgThrHisCysHisArg 107
||||| : : : : :
Db 225 TCATCCAATCTCAGCCGT-----TAACGTCTCTGCTCAAGCTTCCCTCACGCG 272
QY 108 GlyIleGlyArg-----TrpValArgArg-ArgArgArgAsnGlyAlaAlaPr 123
||||| : : : : :
Db 273 TGATTTTCCCGCCCTTTCAGAACTACTGAAAGAGGGAAGAAATAACGGA----- 324
QY 123 oglyGluAlaProHisSerProValIysGluIysProValMetSerAsnIleGlyLysSe 143
||||| : : : : :
Db 325 -----AAAGAG-----AAAGC 335
QY 143 rThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuGlu 163
: ||||| : : : : :
Db 336 ACAGAACATCGTGTGGCAGAGATTGATATGCAGATGCGACAGACACAACTTCTTCA 395
QY 163 yGlnLysGlyCysValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAl 183
||||| : : : : :
Db 396 ACAAAAGGGTTGTGCTTTGGATCAGTGTCTCAGTGTTCAGGGAAGACACTGTTGC 455
QY 183 aCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrrValLeuAspGlyAs 203
||||| : : : : :
Db 456 TTGTGCACCTAAGTAAAGCATTTGTTGAAAGAGGCAAACTTACTTACACACTCGACGGCA 515
QY 203 pAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheIysAlaGluAspArgAlaGluAs 223
||||| : : : : :
Db 516 CAATGTCCGTACGCGCTTAAACCGTACCTCTTCAAGCAGAGACCCGCCGCAAAA 575
QY 223 nIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAl 243
||||| : : : : :
Db 576 CATTAGAGAGANTGTGAGTGGCTAAGTTGTTGCTGACGTCGGAGTCATTTGTATAGC 635
QY 243 aSerLeuIleSerProTyrrArgAspArgAspAlaCysArgAlaLeuLeuProHisSe 263
||||| : : : : :
Db 636 AAGTTTGATTCTCCGTACCGAGAGACAGAGACGCGTCCGCTCTTGTACCTGACGG 695
QY 263 rAsnPheIleGluValPheIleAspLeuProLeuIysIleCysGluAlaArgAspProly 283
||||| : : : : :
Db 696 CGATTTCGTCGAGGTCCTCATGGAGCTTCTTCTATGTGCGAGTCGAGAGATCCAAA 755
QY 283 sGlyLeuTyrrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspPr 303
||||| : : : : :
Db 756 GGGGTTGTACAAGCTTGCACGTGCGAGGCAAAATCAAGGCTTCACTGGNATCGACGCC 815
QY 303 oTyrrGluProProIleAsnGlyGluIleValIleLys---MetLysAspGluGluCysPr 322
||||| : : : : :
Db 816 TTACGAGGCCACCAATTGCGAGGTAGTGTGAAACACACAGGAGACGACGAGTCTGTG 875
QY 322 oSerProLysAlaMetAlaLysGlnValLeuCysTyrrLeuGluGluAsnGlyTyrrLeuGl 342
||||| : : : : :
Db 876 TTCGCCACGTCAGATGGCTGAGAACATCTCTTACTTGCAGAAACAAGGTTATCTTGA 935
```

QY 342 n 342
Db 936 G 936

Search completed: September 9, 2005, 14:24:24
Job time : 4078 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 9, 2005, 13:00:56 ; Search time 531 Seconds
(without alignments)
3823.863 Million cell updates/sec

Title: US-10-829-432-4

Perfect score: 1826

Sequence: 1 RPFHINQTEPLVTHQQPP.....PKMAKQVLCYLENGYLQA 343

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10829432/runat_08092005_132222_25068/app_query.fasta_1.519
-DB=N Geneseq_16Dec04 -OFT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10829432 @CNG_1_1_708 @runat_08092005_132222_25068 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_16Dec04:*

1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002as: *
7: Geneseqn2002bs: *
8: Geneseqn2003as: *
9: Geneseqn2003bs: *
10: Geneseqn2003cs: *
11: Geneseqn2003ds: *
12: Geneseqn2004as: *
13: Geneseqn2004bs: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1826	100.0	1217	AAZ50160	Corn Aden
2	887	48.6	890	AAZ50159	Corn Aden
3	876	48.0	928	AAZ50163	Wheat Ade
4	852	46.7	936	AAZ50162	Soybean A
5	834	45.7	627	AAZ50161	Arabidops

6	817	44.7	1175	3	AAC36588	Aac36588 Arabidops
7	816	44.7	831	6	ABZ12905	Abz12905 Arabidops
8	816	44.7	831	6	ADG87955	Adg87955 A. thalia
9	816	44.7	831	12	ADN74374	Adn74374 Thale cre
10	816	44.7	1068	3	AAC35266	Aac35266 Arabidops
11	816	44.7	1183	3	AAC45659	Aac45659 Arabidops
12	809.5	44.3	714	3	AAC51190	Aac51190 Arabidops
13	788	43.2	687	12	ADJ39869	Adj39869 Plant cDN
14	778.5	42.6	917	3	AAC298337	Aac298337 A. thalia
15	778.5	42.6	948	3	AAC42294	Aac42294 Arabidops
16	757.5	41.5	1130	3	AAC49070	Aac49070 Arabidops
17	752.5	41.2	1133	3	AAC36627	Aac36627 Arabidops
c	640	35.0	378	12	ADJ44283	Adj44283 Plant cDN
19	617.5	33.8	483	3	AAC51315	Aac51315 Arabidops
20	571	31.3	603	8	ACA27590	Act27590 Prokaryot
c	561	30.7	595	12	ADJ43843	Adj43843 Plant cDN
22	555	30.4	600	13	ADT42933	Adt42933 Bacterial
23	534.5	29.3	609	13	ADT42467	Adt42467 Bacterial
24	521.5	28.6	597	8	ACA46934	Act46934 Prokaryot
25	521.5	28.6	615	6	ABN92217	Abn92217 Staphyloc
26	521.5	28.6	615	13	ADS02000	AdS02000 Staphyloc
27	519.5	28.5	588	13	ADT41933	Adt41933 Bacterial
28	517	28.3	681	10	ADF02710	Adf02710 Bacterial
29	516	28.3	1863	13	ADS58748	AdS58748 Bacterial
30	514.5	28.2	591	8	ACA22457	Act22457 Prokaryot
31	514.5	28.2	612	8	ACA44895	Act44895 Prokaryot
32	514	28.1	28401	10	ADC97728	Adc97728 Rhizobito
33	510.5	28.0	1842	13	ADS63529	AdS63529 Bacterial
34	510.5	28.0	1869	13	ADS84272	AdS84272 Bacterial
35	510.5	28.0	1890	13	ADS63901	AdS63901 Bacterial
36	510	27.9	642	8	ACA53672	Act53672 Prokaryot
37	505.5	27.7	597	8	ACA47863	Act47863 Prokaryot
c	503.5	27.6	6715	13	ADR84259	Adr84259 Aspergill
39	503	27.5	603	8	ACA35030	Act35030 Prokaryot
40	503	27.5	651	11	ACH97008	Ach97008 Klebsiell
41	502.5	27.5	2646	6	AAS94855	Aas94855 Human DNA
42	501	27.4	633	10	ACF69692	Acf69692 Photorhab
c	501	27.4	110000	10	ACF67367	26 Continuation (27 o
44	501	27.4	110000	10	ACF65386	2 Continuation (3 o
45	498.5	27.3	1983	13	ADT42849	Adt42849 Bacterial

ALIGNMENTS

RESULT 1
AAZ50160
ID AAZ50160 standard; cDNA; 1217 BP.
AC
XX AAZ50160;
XX
DT 04-MAY-2000 (first entry)
XX
DE Corn Adenyllysulphate kinase-2 cDNA clone.
XX
KW Adenyllysulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
KW 3'-phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW corn; clone p0016.ctscj40rb; transgenic plant; screen; antibody; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
CDS 2..1033
FT /*tag= a
FT /product= "Corn APS kinase-2"
FT /note= "Derived from clone p0016.ctscj40rb"
XX
XX WO200004165-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 98WO-US015809.
XX
PR 14-JUL-1998; 98US-0092833p.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX PI Falco SC, Allen SM, Anderson SL;
 XX XX
 DR WPI; 2000-182430/16.
 DR P-PSDB; AAY44789.
 XX
 PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'
 PT phosphosulfate kinase, useful for altering expression of sulfate
 PT assimilation protein in plants.
 XX
 XX Claim 3; Page 30; 42pp; English.
 XX
 CC The present sequence is a cDNA encoding corn adenylylsulphate kinase (APS
 CC kinase), also known as adenosine-5' phosphosulphate kinase. This is
 CC obtained from clone p0016.ctscj402b, derived from corn pooled tassel
 CC shoots, p0016 cDNA library. APS kinase is a sulphate assimilation
 CC protein, that catalyses the conversion of adenosine-5' phosphosulphate
 CC (APS) to 3'-Phospho-adenosine-5' phosphosulphate (PAPS). This sequence is
 CC used as probe and primers to identify, obtain and synthesise sulphate
 CC assimilation proteins from other plants. It is also used to produce
 CC transgenic plants, that are useful for altering the expression levels of
 CC a sulphate assimilation protein. The APS kinase peptides are useful for
 CC producing antibodies, that are used to screen and isolate cDNA clones
 XX
 SQ Sequence 1217 BP; 344 A; 292 C; 303 G; 278 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.61e-103 Length: 1217
 Score: 1826.00 Matches: 343
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-829-432-4 (1-343) x AAZ50160 (1-1217)

QY 1 ArgProPheHisPheIleAsnGlnThrGluProLeuValThrHisThrGlnGlnProPro 20
 DB 2 CGTCGGTTTCATTTTCATCANTCAACAGAACTCTGGTCACACACAGCAGCACACCG 61
 QY 21 SerProAlaProGlyProAlaSerGlnGlyGlnArgGlnGlnGlnGlnGlnGlnGln 40
 DB 62 AGCCAGCGCCGCCGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 121
 QY 41 ThrProThrLeuAlaValIleValLeuValAsnProGlnArgAlaProProValLeuProGly 60
 DB 122 AGCCAGCGCTCGCGCTCATCTCGTAAATCCACAGCGCGCGCTCCCGTCTCCAGGC 181
 QY 61 LeuThrProSerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrProArgSer 80
 DB 182 CTCACCCCTAGCGATCGCCATCTCCGGCGCTCGTATCTCATGGCTCTACTCCCGTTCC 241
 QY 81 SerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgGluGlyGlyGlyGly 100
 DB 242 TCACACTCTTCGGGGTCTGCCAGTGTAGTGTGGCGCGCGCGAGGGGAGGGCGCGGT 301
 QY 101 AlaArgThrHisCysHisArgGlyIleGlyArgGlyValArgArgArgArgArgArg 120
 DB 302 GCGCGTACGCTGCGCCACCGCGCATTTGGCGGTGGTGGCGCGCGCGCGCGCGGAATGA 361
 QY 121 AlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsnIle 140
 DB 362 GCAGCGCCCGGGAGGCCCGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 421
 QY 141 GlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLys 160
 DB 422 GGGAAATCGACTAAATATTTATGGACAAATTCCTGATGGACAATCTGATAGACAGAAA 481
 QY 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer 180
 DB 482 TTGCTGGGCAAAAGAGGCTGTGTGTATGGATAACAGGACTCAGTGGTTTCAGGGAAAAGT 541

QY 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeu 200
 DB 542 ACTCTTGATGTGCACTGAGTCTGAGTTCGATTCGAGAGGCCACCTCAGCTATGTACTT 601
 QY 201 AspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArg 220
 DB 602 GATGGTGCAACCTCAGACATGGCTAAATAGAGATTTAAGCTTTAAGGCAGAAGACCGT 661
 QY 221 AlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIle 240
 DB 662 GCAGAAATATACGAAGAGTTGGTCAAGTGGCAAGCTTTTGTGTGATGCTGGGTGCATA 721
 QY 241 CysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeu 260
 DB 722 TGCAITGGTACTTGATATCTCCATACAGGAGAGATCGTGATGATGCGGTCTACTT 781
 QY 261 ProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArg 280
 DB 782 CCACATTCTAACTTTATTGAAAGTATTTATTGATTGGCCCTAAAAAATTTGTGAAGCTCGT 841
 QY 281 AspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIle 300
 DB 842 GATCTAAAGGCTTATACAGCTTGCAGTACAGGAAGATTAAGAGGTTTCACTGGAATT 901
 QY 301 AspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGluGlu 320
 DB 902 GATGATCCATACGAAACCAACCAATTAATGTCGAGATAGTAATTAAGATCAAGATGAGGAA 961
 QY 321 CysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyr 340
 DB 962 TGCCCTTCACCCAAAGCAATGCCAAGCAAGTTCTATCTACTCTTGAAGAAAACGATAT 1021
 QY 341 LeuGlnAla 343
 DB 1022 TTGCAAGCT 1030

RESULT 2
 AAZ50159
 ID AAZ50159 standard; cDNA; 890 BP.
 XX
 AC AAZ50159;
 XX
 DT 04-MAY-2000 (first entry)
 XX
 DE Corn Adenylylsulphate kinase-1 cDNA clone.
 XX
 KW Adenylylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
 KW 3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
 KW corn; clone cen3n.pk0088.b10; transgenic plant; screen; antibody; ss.
 XX
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 CDS 3..677
 FT /*tag= a
 FT /product= "Corn APS kinase-1"
 FT /note= "Derived from clone cen3n.pk0088.b10"
 XX
 PN WO200004165-A1.
 XX
 PD 27-JAN-2000.
 XX
 PF 13-JUL-1999; 99WO-US015809.
 XX
 PR 14-JUL-1998; 98US-0092833P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Falco SC, Allen SM, Anderson SL;
 XX
 DR WPI; 2000-182430/16.
 DR P-PSDB; AAY44789.

XX New nucleic acid molecule and chimeric gene encoding an adenosine-5'
PT phosphosulphate kinase, useful for altering expression of sulfate
PT assimilation protein in plants.
XX
XX
PS Claim 3; Page 29; 42pp; English.
PS
XX
XX The present sequence is a cDNA encoding corn adenylylsulphate kinase (APS
CC kinase), also known as adenosine-5' phosphosulphate kinase. This is
CC obtained from clone cen3n.pk0088.b10, derived from corn endosperm, cen3n
CC cDNA library. APS kinase is a sulphate assimilation protein, that
CC catalyses the conversion of adenosine-5' phosphosulphate (APS) to 3'-
CC Phospho-adenosine-5' phosphosulphate (PPS). This sequence is used as
CC probes and primers to identify, obtain and synthesise sulphate
CC assimilation proteins from other plants. It is also used to produce
CC transgenic plants, that are useful for altering the expression levels of
CC a sulphate assimilation protein. The APS kinase peptides are useful for
CC producing antibodies, that are used to screen and isolate cDNA clones
XX
XX Sequence 890 BP; 232 A; 239 C; 262 G; 157 T; 0 U; 0 Other;
SQ

Alignment Scores:	
Pred. No.:	6,58e-46
Score:	887.00
Length:	890
Matches:	171
Percent Similarity:	84.23%
Conservative:	16
Best Local Similarity:	77.03%
Mismatches:	35
Query Match:	48.58%
Indels:	0
DB:	3
Gaps:	0

US-10-829-432-4 (1-343) x AAZ50159 (1-890)

[illegible]

Qy 141 GlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLys 160

Db 66 GGGAAATCGACGAACATCCTGTGGCATGAGTGGCCATCGGGCAGAAAGGACGACAGGT 125

Qy 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyGlyLysSer 180

Db 126 CTGCTGAACCAAGGGCTGCGTCTGTGGATCACTGGCCCTAAGCGGTTACGGGAAAAGC 185

QY 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeu 200

D_b 186 ACGCTCGCGTGC CGCTGAGCCGCGAGCTGCACGGCAGAGGCCACCTCACGTACGTCCTC 245

QY 201 AspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArg 220

D_b 246 GACGGCGACAACCTCAGGCACGGGCTGAACAGGGACCTCAGCTTCGGAGCAGAGGACCGC 305

QY 221 AlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIle 240

Db 306 GCCGAGAACATCCGCAGAGTAGGGGAAGTAGCGAAGCTGTTCCGCCGACGCTGGCCCTCGTC 365

QY 241 CysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeu 260

Db 366 TGCATCGCCAGCCTCATATCGCCCTACAGAGCGACCGGTCGTGCGGATCTGCTG 425

Qy 261 ProHisSerAsn phe Ile Glu Val phe Ile Asp Leu Pro Leu Lys Ile Cys Glu Ala Arg 280

Db 426 CCCAAGCACTCGTTTATCGAGGTGTTCCTGGACGTGCCGCTTCAAGTGTGCGAAGCCAGG 485

Qy 281 AspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIle 300

Db 486 GACCCCAAGGCCTCTACAAGCTCGCACGGCGCGCAAAATCAAAGGTTCAACGGCATC 545

Qy 301 AspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGluGlu 320

Db 546 GACGATCCTTACGAACCGCCGTCGGACTGTGAGATAGTGATCCAGTGTAAGTCGGCGAC 605

Qy 321 CysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyr 340

[illegible]

Db	606	TGCCCTTCGCCCTGAATCGATCGCTGGTCA	CGTTGTCGTACCTTGAGACGAATCGTTTC	665
Qy	341	LeuGln	342	
Db	666	CTCCAG	671	
RESULT 3				
AAZ50163				
ID	AAZ50163	standard; cDNA; 928 BP.		
XX				
AC	AAZ50163;			
XX				
DT	04-MAY-2000	(first entry)		
XX				
DE	Wheat Adenylylsulphate kinase-1	CDNA clone.		
XX				
KW	Adenylylsulphate kinase; Adenosine-5'	phosphosulphate kinase; APS kinase;		
KW	3'-Phospho-adenosine-5'	phosphosulphate; PAPS; sulphate assimilation;		
KW	wheat; clone wrl.pk0101.e2;	transgenic plant; screen; antibody; ss.		
XX				
OS	Triticum aestivum.			
XX				
PH	Key	Location/Qualifiers		
FT	CDS	3..743		
FT		/*tag= a		
FT		/product= "Wheat APS kinase-1"		
FT		/note= "Derived from clone wrl.pk0101.e2"		
FT				

WO200004165-A1.

27-JAN-2000.

13-JUL-1999; 99WO-US015809.

14-JUL-1998; 98US-0092833P.

(DUPO) DU PONT DE NEMOURS &

Falco SC, Allen SM, Anderson SL;

WPI; 2000-182430/16.

100-100000-1

new nucleic acid molecule and enzyme gene encoding an adenosine 5'-phosphosulfate kinase, useful for altering expression of sulfate assimilation protein in plants.

Claim 3; Page 34-35; 42pp; English.

The present sequence is a cDNA encoding wheat adenylylsulphate kinase (APS kinase), also known as adenosine-5' phosphosulphate kinase. This is obtained from clone wri.pk0101.e2, derived from 7 day old light grown wheat root seedlings, wri cDNA library. APS kinase is a sulphate assimilation protein, that catalyses the conversion of adenosine-5' phosphosulphate (APS) to 3'-phospho-adenosine-5' phosphosulphate (PAPS). This sequence is used as a probe and primer to identify, obtain and synthesise sulphate assimilation proteins from other plants. It is also used to produce transgenic plants, that are useful for altering the expression levels of a sulphate assimilation protein. The APS kinase peptides are useful for producing antibodies, that are used to screen and isolate cDNA clones

Sequence 928 BP: 285 A: 168 C: 250 G: 225 T: 0 U: 0 Other:

Alignment Scores:

Pred. No.:	3.27e-45	Length:	928
------------	----------	---------	-----

Score:	876.00	Matches:	174
--------	--------	----------	-----

Percent Similarity:	Conservative:
79.37%	26

Query Match:	47.97%	Indels:	9
--------------	--------	---------	---

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525


```
Db 460 CTGAGTCAAAAGCTTGCACTCCAAAGGAAACTGTCTTATCATCTCTGATGTCACATATT 519
Qy 206 ArgHisGlyLeuAenArgAspLeuSerPheLeuAlaGluAspArgAlaGluAenIleArg 225
Db 520 CGGCATGGTCTAAACAGAGATCTTAGTTTGTAGAGCAGAGATCGTTCTGAAACATTAGA 579
Qy 226 ArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeu 245
Db 580 AGGATTGGTGAGGTGGCAAACTCTTTGCAGATGCTGGTGTATTGTCATCATAGTTTA 639
Qy 246 IleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSerAsnPhe 265
Db 640 ATATACCATTACCAAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 699
Qy 266 IleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeu 285
Db 700 ATTGAGTTTTCATAGATGTTCCACTACATGTTGTGAAGCTAGGACCCAAAGGACATC 759
Qy 286 TyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspProTyrGlu 305
Db 760 TACAAGCTTGCTCGAGCTGGAAGATCAAAAGGTTTCACTGGTATAGATCATCATATGAA 819
Qy 306 ProProIleAsnGlyGluIleValIleLysMetLysAspGluGluCysProSerProLys 325
Db 820 CCACCGTGTAGTGTGATAGATAGTATTACAAACAGAAAGGAGTCACTGTAAAGTCTCCAGT 879
Qy 326 AlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuGlnAla 343
Db 880 GATATGGCTGAAGAGTATATCTTACTTGGAGGAGACGGATACCTGCGGGCT 933

RESULT 5
AAC44691
ID AAC44691 standard; DNA; 627 BP.
XX
AC AAC44691;
XX
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 43768.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 28-APR-1999; 99US-0130510P.
PR 30-APR-1999; 99US-0130891P.
PR 04-MAY-1999; 99US-0131449P.
PR 05-MAY-1999; 99US-0132048P.
PR 06-MAY-1999; 99US-0132407P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
```

```
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148315P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.

PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 8.2e-43
Score: 834.00
Percent Similarity: 88.35%
Best Local Similarity: 74.76%
Query Match: 45.67%
DB: 3
Gaps: 0

US-10-829-432-4 (1-343) x AAC44691 (1-627)

Qy 137 MetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyInSer 156
Db 1 ATGTCGACAGTGGGAAATTCACGACATATTTGGCAAGAAATCCCCCATTTGGGAAACT 60

Qy 157 AspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGly 176
Db 61 GAAAGGCAGAAAGTTGTTAAATCAGAAAGGTTGTGTGTGTGGATCACAGGGCTCAGTGGC 120

Qy 177 SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu 196
Db 121 TCAGGGAAACACGCTAGCTTGCTTCGTATAGAGAGCTGAAACACCGGGAAAGCTA 180

Qy 197 ThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLys 216
Db 181 TCATATATTTGATGGTGACAATCTTCGTCTCATGTTTGAACAAAGATCTTGTGTTCAAG 240

Qy 217 AlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAsp 236
Db 241 GCAGAGGATAGATGGGAAATATACGAGGTCGAGAGAGTAGCCAAACTCTTTGCGGAT 300

Qy 237 AlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAspAlaCys 256
Db 301 GCTGGTTTAACTGTATTGTCAGGCTCATATCCCGTATAGAAAGACCGTGACGCTGC 360

Qy 257 ArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIle 276
Db 361 AGGGAATGATACAGAATTCATCTTTTATTTCAGGTTTTCATGAACATGCTCTGCAATTG 420

Qy 277 CysGluAlaArgAspProLysGlyLeuTyrIlyLeuAlaArgThrGlyIleLysGly 296
Db 421 TGTGAAGCAAGGAGCCCTAAAGGCTATACAAAGCTTGACGTGCGAGGAAGATCAAGGT 480

Qy 297 PheThrGlyIleAspAspProTyrGluProPheIleAsnGlyGluIleValIleLysMet 316
Db 481 TTCACAGGAATAGATGATCCGTATGAGTCTCCCTTGAACCTGTGATAGATTGAAAGAG 540

Qy 317 LysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGlu 336
Db 541 AAAGAGGAGAGAGTGTCTTCCCTGTAGCTATGCTGAGGAAGTGTCTCTTATCTAGAA 600

Qy 337 GluAsnGlyTyrLeuGln 342
Db 601 GACAAAGGTTTCTCTCAA 618

RESULT 6
AAC36588
ID AAC36588 standard; DNA; 1175 BP.
XX
AC AAC36588;
```

XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 14353.
DE
DE
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
XX 09-MAR-1999; 99US-0123348P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
XX 23-MAR-1999; 99US-0125788P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
XX 29-MAR-1999; 99US-0126264P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
XX 01-APR-1999; 99US-0126785P.
PR 29-JUN-1999; 99US-0140991P.
XX 06-APR-1999; 99US-0128234P.
PR 30-JUN-1999; 99US-0141287P.
XX 08-APR-1999; 99US-0128714P.
PR 01-JUL-1999; 99US-0141842P.
XX 19-APR-1999; 99US-0130077P.
PR 02-JUL-1999; 99US-0142055P.
XX 21-APR-1999; 99US-0130449P.
PR 06-JUL-1999; 99US-0142390P.
XX 23-APR-1999; 99US-0130891P.
PR 08-JUL-1999; 99US-0142803P.
XX 28-APR-1999; 99US-0131449P.
PR 09-JUL-1999; 99US-0142920P.
XX 30-APR-1999; 99US-0132407P.
PR 12-JUL-1999; 99US-0142977P.
XX 04-MAY-1999; 99US-0132484P.
PR 13-JUL-1999; 99US-0143542P.
XX 05-MAY-1999; 99US-0132485P.
PR 14-JUL-1999; 99US-0143624P.
XX 06-MAY-1999; 99US-0132486P.
PR 15-JUL-1999; 99US-0144005P.
XX 07-MAY-1999; 99US-0132863P.
PR 16-JUL-1999; 99US-0144085P.
XX 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138647P.
PR 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149722P.
XX 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
XX 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.

XX
DR
XX
PT
PT
XX
XX
PS
XX

WPI; 2002-304127/34.

Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.

Claim 144; SEQ ID NO 710; 577pp + Sequence Listing; English.

The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

SQ Sequence 831 BP; 224 A; 137 C; 223 G; 247 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.42e-41 Length: 831
Score: 816.00 Matches: 159
Percent Similarity: 77.87% Conservative: 31
Best Local Similarity: 65.16% Mismatches: 44
Query Match: 44.69% Indels: 10
DB: 6 Gaps: 3

US-10-829-432-4 (1-343) x AB212905 (1-831)

QY 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrrValArgArgArgArgAsn 119
DB 127 GGATCTCAAACTGTAGTCATACAAATAATGGATCTATTCCTGAGGTAAATCCATTAAC 186
QY 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
DB 187 GGT-----CACAGGACAAAGACAGACCT---TTGCTACG 222
QY 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
DB 223 GTCGAAACTCGCAAAATATAAGTGCATGAATGTTCTGTTGAGAAAGTTGATGACAG 282
QY 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLysSerGlySerGlyLys 179
DB 283 AGATTGCTTGATCAGAAAGGATGTGTGATTTGGTTCACCGGCTTAGTGTTTCAGGGAAG 342
QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
DB 343 AGTACTTGCGTGTGCTTGTGAATCAGATGTTGTATCAAAAGGGAAGCTTTGTATATT 402
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
DB 403 CTTGATGGTGATAATGTTAGGCATGCTTAAACCGTGATCTTAGCTTTAAAGCTGAGAT 462
QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
DB 463 CGTGCAAGAAATATTCGTAGAGTTGGAGAGGTTGCTTAAGCTTTTTCGGATGCTCGAATA 522
QY 240 IleCysIleAlaSerLeuIleSerProTrpArgArgAspAlaCysArgAlaLeu 259
DB 523 ATCTGATTCGCGAGTTGATATCTCTTATAGAACAGATAGGAGCCTTGTGGAAGTTG 582
QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
DB 583 CTCCTCCGAGGAGATTTGTTGAGGTGTTTCATGATGTACCGCTTAGTGTTCGAGGCG 642
QY 280 ArgAspProLysGlyLeuTrpLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
DB 643 AGGGATCCAAAGGGCTTTTACAAGCTTGCTGTCGAGGAAAGATCAAAAGGTTTTCACGGG 702

QY 300 IleAspAspProTrpGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
DB 703 ATCGATGACCTTACGAGCCACCAATTGAACCTCGAG-----ATTCTCTAGGACGTGAA 756
QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTrpLeuGluGluAsnGly 339
DB 757 GGAGAACTTCTCTATCGAAATGCGGAAGAGTTCGCGATACTTAGATAACAAGGCT 816
QY 340 TyrLeuGlnAla 343
DB 817 TATCTTCAAGCA 828
RESULT 8
ADG87955
ID ADG87955 standard; cDNA; 831 BP.
XX
AC ADG87955;
XX
DT 22-APR-2004 (first entry)
XX
DE A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #397.
XX
KW Pathogen infection-related gene; plant; Peronospora parasitica;
KW defence mechanism; RPP7; RPP8; pathogen resistance; transgenic plant;
KW oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS Arabidopsis thaliana.
XX
PN WO200222675-A2.
XX
PD 21-MAR-2002.
XX
PF 14-SEP-2001; 2001WO-US028506.
XX
PR 15-SEP-2000; 2000US-0232778P.
PR 22-JUN-2001; 2001US-0300183P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
PA (UYNC-) UNIV NORTH CAROLINA.
PA (GLAZ/) GLAZEBROOK J.
PA (WANG/) WANG X.
PA (DANG/) DANG J L.
PA (EULG/) EULGEM T.
PA (ZHUT/) ZHU T.
XX
PI Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;
XX
XX WPI; 2002-292409/33.
XX
PT Novel isolated polynucleotide, useful for conveying pathogen resistance
PT to plants, and for identifying plants infected with a pathogen.
XX
PS Claim 3; SEQ ID NO 397; 605pp; English.
XX
CC The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC ADG87557) whose expression is altered in response to pathogen infection,
CC and to homologues of these genes from other plants or fungi, especially
CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC expression of genes of the invention was upregulated or downregulated in
CC Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC indicating that they play a role in defence mechanisms. The genes of the
CC invention are regulated by RPP7 or RPP8 which act via unconventional
CC signalling cascades, or by the RPP4-dependent pathway. The invention also
CC relates to polypeptides encoded by the pathogen infection-related genes;
CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC ; expression cassettes, host cells and pathogen-resistant transgenic
CC plants and their progeny comprising a polynucleotide of the invention;
CC and a method of identifying a plant cell infected with a pathogen. The
CC polynucleotide sequences and methods of the invention are useful for
CC identifying plants infected with a pathogen, and for conferring
CC resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC nematodes and insects (e.g., aphids). The present sequence represents an

CC Arabidopsis thaliana gene whose expression is altered in response to
CC Peronospora parasitica infection. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 831 BP; 224 A; 137 C; 223 G; 247 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,42e-41 Length: 831
Score: 816.00 Matches: 159
Percent Similarity: 77.8% Conservativeness: 31
Best Local Similarity: 65.16% Mismatches: 44
Query Match: 44.69% Indels: 10
DB: 6 Gaps: 3

US-10-829-432-4 (1-343) x ADG87955 (1-831)

QY 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsn 119
DB 127 GGATCTCAAACTCTGAGTCATAACAAAATGGATCTATTCTCGAGGTAAATCCATTAC 186
QY 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
DB 187 GGT-----CACACGGGACAAAAGCAAGACCT---TTGTCTAGC 222
QY 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
DB 223 GTCGGAACCTCAGCAATATAAAGTGGCATGAATGTCGTGAGAAAGTGTATGACAG 282
QY 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
DB 283 AGATTGCTTGATCAGAAAGGATGTGTGATTGGGTCCACCGGTCTTAGTGTTCAGGGA 342
QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
DB 343 AGTACTTTGGCTTGCTTTGAATCAGATGTTGTATCAAAAGGGAAGCTTTGTTATATT 402
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelysAlaGluAsp 219
DB 403 CTTGATGGTGATTAATGTAGGATGCTTAACCGTGATCTTAGCTTTAAAGCTGAGGAT 462
QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
DB 463 CGTGACAGAAATATTCGTAGAGTTGGAGAGTTGCTTAAGCTTTTTCGGATCTGGAATA 522
QY 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspAspAlaCysArgAlaLeu 259
DB 523 ATCTGCATTGCGAGTTTGATATCTCTCTATAGACAGATAGGACGCTTTGTCGAAGTTTG 582
QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
DB 583 CTCCTCCGAGGAGATTTTGTAGGTGTCATGGAATGACCCCTAGTGTTCGAGGCG 642
QY 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
DB 643 AGGGATCCAAAGGCTTTTACAAGCTTGCTGTCGAGGAAAGATCAAGGTTTTCACGGG 702
QY 300 IleAspAspProTyrGluProPheIleAsnGlyGluIleValIleLysMetLysAspGlu 319
DB 703 ATCGATGACCCCTTACGAGCCACCATGAACTCGAG-----ATTCTCTAGAGCGTAA 756
QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrIleGluGluAsnGly 339
DB 757 GGAGGAACTTCTCTATCGAATGGCGGAAGAGGTCTCGGATACTTAGATAACAGGCT 816
QY 340 TyrLeuGlnAla 343
DB 817 TATCTTCAAGCA 828

RESULT 9

ADN74374

ID ADN74374 standard; cDNA; 831 BP.

XX AC ADN74374;
XX DT 15-JUL-2004 (first entry)
XX DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2269.
XX KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;
XX KW growth regulator; animal feed product; thale cress;
XX KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX OS Arabidopsis thaliana.
XX PN WO2004035798-A2.
XX PD 29-APR-2004.
XX PF 20-OCT-2003; 2003WO-EP011658.
XX PR 18-OCT-2002; 2002EP-00079408.
XX PA (CROP-) CROPDESIGN NV.
XX PI Inze D, De Veylder L, Vlieghe K;
XX DR WPI; 2004-348466/32.
XX DR P-PSDB; ADN74375.
XX PT Altering plant characteristics, useful for producing plants for enzyme or
XX PT pharmaceutical production comprises modifying in a plant, expression of
XX PT one or more nucleic acids and/or modifying level or activity of one or
XX PS Claim 1; SEQ ID NO 2269; 134pp; English.
XX CC This invention relates to a novel method for altering one or more plant
XX CC characteristics. Specifically, it refers to identifying genes that are up
XX CC - or down-regulated in transgenic plants overexpressing the heterodimeric
XX CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
XX CC alter plant characteristics accordingly. The present invention describes
XX CC generating transgenic plants for the production of growth regulators,
XX CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
XX CC the altered plant characteristics are selected from increased yield or
XX CC biomass, enhanced survival capacity, stress tolerance, plant architecture
XX CC or physiology, altered endoreplication, biochemistry, signal
XX CC transduction, storage lipid mobilisation and/or altered photosynthesis,
XX CC each relative to the corresponding wild type plants. Accordingly, these
XX CC sequences can also be useful as positive or negative selectable markers
XX CC during transformation of cells or tissues. The identified genes play a
XX CC role in a variety of biological processes such as DNA replication, cell
XX CC wall biosynthesis, nitrogen and/ or carbon metabolism or they function as
XX CC transcription factors. This polynucleotide sequence is thale cress cDNA
XX CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa
XX CC transcription factor, given in an exemplification of the invention.
SQ Sequence 831 BP; 224 A; 137 C; 223 G; 247 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,42e-41 Length: 831
Score: 816.00 Matches: 159
Percent Similarity: 77.8% Conservativeness: 31
Best Local Similarity: 65.16% Mismatches: 44
Query Match: 44.69% Indels: 10
DB: 12 Gaps: 3

US-10-829-432-4 (1-343) x ADN74374 (1-831)

QY 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsn 119
DB 127 GGATCTCAAACTCTGAGTCATAACAAAATGGATCTATTCTCGAGGTAAATCCATTAC 186
QY 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139

```
Db 187 GGT-----CACACGGGACAAAAGCAAGGACCT---TTGTCTACG 222
QY 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyClnSerAspArgGln 159
Db 223 GTCGGAACTCGACAATAATAAGTGGCATGAATGTTCTGTGAGAAAAGTTGATAGACAG 282
QY 160 LysLeuLeuGlyClnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db 283 AGATTGCTTGATCAGAAAGATGTGTGATTGGTCCACCGTCTTAGTGTTCAGGGAAG 342
QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
Db 343 AGTACTTGGCTGTGCTTGAATCAGATGCTGTATCAAAAGGGAAGCTTTGTTATATT 402
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
Db 403 CTTGATGGTGATGAATGTTAGGCATGGCTTAACCGTGATCTTAGCTTTAAAGCTGAGGAT 462
QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 463 CGTGCAGAGAATATTCTGAGAGTGGAGAGGTTGCTAAGCTTTTGGCGATGCTGGAATA 522
QY 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
Db 523 ATCTGCATTGCCAGTTTGTATATCTCTTATAGAACAGATAGGACGCTTGTGCAAGTTTG 582
QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 583 CTCCCGGAGGAGATTTGTTGAGGTTCATGGATGATACCGCTTAGTGTTCGAGGCG 642
QY 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db 643 AGGGATCAAAAGGGTCTTTACAAGCTTGCTCGTCAGGAAAGATCAAAAGGTTTACCGGG 702
QY 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db 703 ATCGATGACCCCTTACGAGCCACCATGAACTCGAG-----ATTCTCTAGGACGTGAA 756
QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 757 GGAGGAACCTTCTTCATCGAATGCGGNAAGGTCGCGGATACTTTAGATNAACAAGGTT 816
QY 340 TyrLeuGlnAla 343
Db 817 TATCTTCAAGCA 828
RESULT 10
AAC35266
ID AAC35266 standard; DNA; 1068 BP.
XX AAC35266;
XX AAC35266;
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9570.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX EP1033405-A2.
PD 06-SEP-2000.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
```



```
Db      676  CTCCCCGAGGAGATTGTTGAGGTGTTTCATGATGATACCGCTTAGTCTTTGCGAGGCG 735
Qy      280  AtgAspProlysglyLeuTyrlyLeuAlaAqgThrGlylylelysglypethrGly 299
Db      736  AGGATCCAAAGGGTCTTTACAGCTTGCTCGTCGAGGAAAGATCAAGGTTTACCCGG 795
Qy      300  ileAspAspProTyrGluProProIleAsnGlyGluileValileLysMetLysAspGlu 319
Db      796  ATCGATGACCTTACGAGCCACCATTTGAATCGGAG-----ATTCTCTAGGACGTGNA 849
Qy      320  GluCysProSeuProLysAlaMetAlaLysGlnValLeuCyTyrLeuGluGluAsnGly 339
Db      850  GGAGGAACCTTCTCTATCGAAATGCGGAAAGGTCGTCGATACTTAGATAACAAGGCT 909
Qy      340  TyrLeuGlnAla 343
Db      910  TATCTTCAAGCA 921

RESULT 11
AAC45659
ID AAC45659 standard; DNA; 1183 BP.
XX
AC AAC45659;
XX
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 47294.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 19-APR-1999; 99US-0128714P.
XX 18-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 21-JUN-1999; 99US-0139763P.
XX 22-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 19-JUL-1999; 99US-0144335P.
XX 20-JUL-1999; 99US-0144352P.
XX 20-JUL-1999; 99US-0144632P.
XX 20-JUL-1999; 99US-0144884P.
XX 21-JUL-1999; 99US-0144814P.
XX 21-JUL-1999; 99US-0145086P.
XX 22-JUL-1999; 99US-0145088P.
XX 22-JUL-1999; 99US-0145085P.
XX 22-JUL-1999; 99US-0145087P.
XX 22-JUL-1999; 99US-0145089P.
XX 22-JUL-1999; 99US-0145192P.
XX 23-JUL-1999; 99US-0145145P.
XX 23-JUL-1999; 99US-0145218P.
XX 23-JUL-1999; 99US-0145224P.
XX 26-JUL-1999; 99US-0145276P.
XX 27-JUL-1999; 99US-0145913P.
XX 27-JUL-1999; 99US-0145918P.
XX 27-JUL-1999; 99US-0145919P.
XX 28-JUL-1999; 99US-0145951P.
XX 02-AUG-1999; 99US-0146386P.
XX 02-AUG-1999; 99US-0146388P.
XX 02-AUG-1999; 99US-0146389P.
```

```
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-014902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.

PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-014902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.

PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 2.08e-41 Length: 1183
Score: 816.00 Matches: 159
Percent Similarity: 77.87% Conservative: 31
Best Local Similarity: 65.16% Mismatches: 44
Query Match: 44.69% Indels: 10
DB: 3 Gaps: 3

US-10-829-432-4 (1-343) x AAC45659 (1-1183)
QY 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgAsn 119
Db 221 GGATCTCAAACTCGAGTCATAACAAAATGATCTATTCTCGAGGTAAATCCATTAC 280
QY 120 GlyAlaAlaProGlyGluAlaProHisSerProValysGluLysProValMetSerAsn 139
Db 281 GGT-----CACACGGGACAAAAGCAAGGACCT---TTGTCTAGC 316
QY 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db 317 CTCGAAACTCGACAAATATAAAGTGGCATGAATGTTCTGTGAGAAAAGTTGATAGACAG 376
QY 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db 377 AGATTGCTTGATCAGAAAGGATGTGATTTGGGTACCCTCTTAGTGGTTCAGGGAAG 436
QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
Db 437 AGTACTTTGGCTTGCTTTGAATCAGATGTTATCAAAAGGGGAAGCTTTGTATATT 496
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
Db 497 CTTCATGTTGATAATGTTAGGATGGCTTAACCGTGATCTTAGCTTTTAAAGCTGAGGAT 556
QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 557 CGTCAGAGATATTCGTAGAGTTGGAGAGTTGCTTAAGCTTTTTCGGATGCTGAATA 616
QY 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
Db 617 ATCTGCATTGCGAGTTTGATATCTCTTATAGAACAGATAGGACGCTTGTGGAAGTTG 676
QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 677 CTCCCCGAGGAGATTTTGTAGGTGTTCAGGATGTTACCGCTTAGTGTTCGAGGCG 736
QY 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db 737 AGGATCCAAAGGGTCTTTACAAGCTTGCTCGCAGGAAAGATCAAAAGTTTACCGGG 796
QY 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db 797 ATCGATGACCTTACGAGCCACCATTAAGTCCGAG-----ATTTCCTAGGACGTGAA 850
QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 851 GGAGGAACCTTCTTCATCGAAATGCGGAAAGGTGTCGGATACTTAGATAACAAGGCT 910
QY 340 TyrLeuGlnAla 343
Db 911 TATCTTCAAGCA 922

RESULT 12
AAC51190
ID AAC51190 standard; DNA; 714 BP.
XX
AC AAC51190;
XX
DT 18-OCT-2000 (first entry)
XX
```


US-10-829-432-4 (1-343) x AAC51190 (1-714)

XX	Eukaryota.
OS	
XX	
XX	US2004016025-A1.
EN	
XX	
XX	26-SEP-2002; 2002US-00260238.
XX	
XX	26-SEP-2001; 2001US-0325277P.
PR	
PR	26-SEP-2001; 2001US-0325448P.
PR	
PR	04-APR-2002; 2002US-0370620P.
XX	
XX	(BUDW/) BUDWORTH P.
PA	
PA	(MOUG/) MOUGHAMER T.
PA	(BRIG/) BRIGGS S P.
PA	(COOP/) COOPER B.
PA	(GLAZ/) GLAZEBROOK J.
PA	(GOFF/) GOFF S A.
PA	(KATA/) KATAGIRI F.
PA	(KREP/) KREPS J.
PA	(PROV/) PROVANT N.
PA	(RICK/) RICK D.

PA (ZHUTU/) ZHU T.
XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI Goff SA, Katagiri F, Kreps J, Provart N, Rieke D, Zhu T;
XX WPI; 2004-190374/18.
XX New rice promoter, useful for manipulating crop plants to alter or
PT improve phenotypic characteristics, e.g. produce large quantities of oil
PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT or high nutritional value.
XX Claim 72; SEQ ID NO 869; 230pp; English.
XX The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC or constitutive transcription of an operatively linked nucleic acid
CC segment. The invention also relates to a method for augmenting a plant
CC genome and a method of identifying a gene, where its expression is
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC encode are useful for manipulating crop plants to alter or improve
CC phenotypic characteristics, to produce large quantities of oil or
CC proteins, to incur resistance to insecticides, viruses or fungi, and to
CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 687 BP; 195 A; 154 C; 165 G; 173 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6,11e-40 Length: 687
Score: 788.00 Matches: 158
Percent Similarity: 83.89% Conservativity: 19
Best Local Similarity: 74.88% Mismatches: 21
Query Match: 43.15% Indels: 13
DB: 12 Gaps: 2

US-10-829-432-4 (1-343) x ADJ39869 (1-687)

QY 138 SerAenIleGlyLysSerThrAenIleLeuTrpHisAenCysLeuIleGlyClnSerAep 157
DB 10 TCAACTGTCGGTAAATCAACAACATCCGATGCGCATGACTGCCCAAGTAAACCCAGCTTGAC 69
QY 158 ArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGly--- 176
DB 70 AGACAAAGTCTGCTGACCAAGAGGGCTGTGTGTTGGATCACTGGCTTAAGTGTTCA 129
QY 177 -----SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgG1 194
DB 130 GTCAACTTCAGGAAAAAGCACCTTGTGATGTGCTGCTGAGCCGTGAGCTCAAGAGG 189
QY 194 yHisLeuThrTrpValLeuAspGlyAspAenLeuArgHisGlyLeuAenArgAspLeuSe 214
DB 190 GCATCTGACCTATGTTCTTGACGGCGACAAATCTCCGGCAGCCCTGGAACAAGATCTCAG 249
QY 214 rPhelysAlaGluAspArgAlaGluAenIleArgArgVal-GlyGluValAlaLysLeuP 234
DB 250 CTTCAAGCCAGGATGCTGCGCAAAATATACGAGAGTTGGAGAGTGGCAAGCTGT 309
QY 234 heAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgA 254
DB 310 TTGCAGATGCTGGATTGATCTGCATTACTACTAGTTTGTATATCACCCCTATAAGAGTGATCGAA 369
QY 254 sPAlaCysArgAlaLeuLeuProHisSerAenPhelleGluValPhelleAspLeuProL 274
DB 370 GCGCGCTGCGCAAAATTACTACCAATCTTCGTTTCAATGAGGTGTTCTGTAATGTCAC 429

QY 274 euLysIleCysGluAlaAArgAspProLysGlyLeuTyrLysLeuAlaAArgThrGlyLysI 294
DB 430 TTGAAGTATGTGAAGAAAGGATCCAAAGGCTGTACAGCTTCTGTCGCGCAAAA 489
QY 234 le-LysGlyPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluLeuVal 313
DB 490 TCAAAGGCTTTACGGGAATAGATGATCTTATGAAACACCTTCAGATTGTGAG----- 543
QY 314 lIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCys 333
DB 544 -----GGGACTGCCCTTCACCTAAATCAATGGCTGATCAAGTAGTGTC 588
QY 334 TyrLeuGluGluAenGlyTyrLeuGln 342
DB 589 TATCTTGAAGCAATGGATTCTTTTCAG 615
RESULT 14
AAZ98337
ID AAZ98337 standard; DNA; 917 BP.
XX
AC AAZ98337;
XX
DT 14-JUN-2000 (first entry)
XX
DE A. thaliana gene involved in environmental stress tolerance.
XX
KW Environmental stress; plant; transgenic plant; anaerobic; flooding; cold;
KW dehydration; drought; heat stress; salinity; osmotolerance; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200008187-A2.
XX
PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-EP005652.
XX
PR 04-AUG-1998; 98EP-00202634.
XX
PA (VLAA-) VLAAIS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI Lee JH, Verbruggen N;
XX
DR WPI; 2000-205726/18.
DR P-PSDB: AAY77957.
XX
PT Isolation of polynucleic acids useful for producing transgenic plant by
PT isolating genes involved in tolerance to environmental stress.
XX
PS Claim 4; Page 218-220; 312pp; English.
XX
CC The invention relates to isolation of coding sequences and/or genes
CC involved in tolerance to environmental stress in plants. The sequences
CC (AAZ98305-298365) are useful for producing a transgenic plant having
CC enhanced tolerance or resistance to environmental stress conditions such
CC as anaerobic, flooding, cold, dehydration, drought, heat stress or
CC salinity. This is useful for producing improved yield, growth,
CC development and productivity under environmental stress conditions, and
CC also provides growth of crops in areas where they cannot grow without the
CC induced osmotolerance. Sequences AAZ98305-365 represent polynucleotide
CC sequences from A. thaliana that are involved in environmental stress
CC tolerance
XX
SQ Sequence 917 BP; 244 A; 220 C; 229 G; 224 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.21e-39 Length: 917
Score: 778.50 Matches: 175
Percent Similarity: 66.04% Conservativity: 37
Best Local Similarity: 54.52% Mismatches: 67
Query Match: 42.63% Indels: 43
DB: 3 Gaps: 9

US-10-829-432-4 (1-343) x AAZ98337 (1-917)

QY 40 ProThrProThrLeuAlaValIleLeuValAnProGlnArgAlaProProVal----- 57
DB 14 CCGTCGGATTCAATGAAGATTAGTATCAGAGCATCTCGA---CCGTGGTTCGTGT 70
QY 58 ---LeuProGlyLeu-----ThrProSerAspAla-----Pro 67
DB 71 TCTATTCAGGTCTCGCGCGGATCCACCCGAAACCTCCAGTGCAGGTTTCTCTCAAG 130
QY 68 LeuProAlaLeuValIleHisGlyLeuThrProArgSerSerHisSerSerAlaGlyLeu 87
DB 131 CTGCTGGGTC-GTCTATTCCGGCGGACAGCCGAAATAGTAGCAATTTCTACTCTCT 189
QY 88 AlaSerAspSerGlyArgGluGlyGluGlyArgGlyAlaArgThrHisCysHisArg 107
DB 190 TCATCCAATCTCAGCCGT-----TAACTCTCTGCTCAAGCTTCCTCACCGC 237
QY 108 GlyIleGlyArg-----TrpValArgArg-ArgArgArgAsnGlyAlaAlaPr 123
DB 238 TGATTTCCCGCCCTTTCAGAACTATACTGAAGAGGGAAGAAATAACGGA----- 289
QY 123 oGlyGluAlaProHisSerProValLysGluLysProValMetSerAsnIleGlyLysSe 143
DB 290 -----AAAGAG-----AAAGC 300
QY 143 rThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuG1 163
DB 301 AGAGAACATCGTGTGGCAGAGATTCGATATGCAGATCGCACAGCAACACTTCTTCA 360
QY 163 yGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAl 183
DB 361 ACAAAAGGGTGTGTCTGTCTGATCTAGTCTCTAGTGTTCAGGGAAGACACTGTTC 420
QY 183 aCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyValLeuAspGlyAs 203
DB 421 TTGTGCACCTAAGTAAAGCATTTGTTGAAAGAGGCAAACTTACTTACACACTCGACGCGA 480
QY 203 pAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAs 223
DB 481 CAATTCCTGTCACGGCCCTTAAACCGTACCTTCAAGCAGAGCACCACCGAAAA 540
QY 223 nIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAl 243
DB 541 CATTAGAAGATTGTGTAGGTGGCTAAGTTGTTCTGACGTCGAGTCAITTTGTATAGC 600
QY 243 aSerLeuIleSerProTyArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSe 263
DB 601 AAGTTTGATTTCTCGGTACCGGAGAGACAGACGCGTCCGCGTCTTGTGTACCTGACGG 660
QY 263 rAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLy 283
DB 661 CGATTTCGTGAGGCTTCATGAGAGTTCCTCTTCATGTGTCGAGTGCAGAGATCCAA 720
QY 283 sGlyLeuTyLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspPr 303
DB 721 GGGGTTGTACAAGCTTGACGTCGAGGCAAAATCAAGGCTTCACTGGAATCGACGACC 780
QY 303 oTyArgLeuProPheIleAsnGlyGluIleValIleLys---MetLysAspGluGluCysPr 322
DB 781 TTACAGGACACAGTGAATTGCGAGGTAGTGTGTAACACACAGGAGACGAGTCTGTG 840
QY 322 oSerProLysAlaMetAlaLysGlnValLeuCysTyLysLeuGluAlaAsnGlyTyLysLeuG1 342
DB 841 TTCGCCAGCTCAGTGGCTGAGAAACATCATCTTACTTGCAGAAAGGTTATCTTGA 900
QY 342 n 342
DB 901 G 901

RESULT 15

AAC42294

ID AAC42294 standard; DNA; 948 BP.

XX AAC42294;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 35008.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 23-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 9, 2005, 13:05:12 ; Search time 198 Seconds
(without alignments)
2834.560 Million cell updates/sec

Title: US-10-829-432-4
Perfect score: 1826
Sequence: 1 RPFHINOTEPVLTHTQPP.....PRMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-Q=/cgn2_1/USFTO_spool_p/US10829432/runat_08092005_132224_25102/app_query.fasta_1.519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10829432@cgn1.1.105@runat_08092005_132224_25102 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	521.5	28.6	615	3	US-09-134-001C-1680
2	517	28.3	681	4	Sequence 1680, Ap
3	503	27.5	651	4	Sequence 2995, Ap
4	496	27.2	2014	4	Sequence 2803, Ap
5	495	27.2	2506	1	Sequence 1, Appl
6	495	27.1	1875	4	Sequence 4, Appl
7	494.5	27.1	2015	4	Sequence 1468, Ap
8	486.5	26.6	2617	4	Sequence 1925, Ap
9	484.5	26.5	1851	4	Sequence 21, Appl
10	484.5	26.5	2000	4	Sequence 10, Appl
11	482	26.4	1845	4	Sequence 2, Appl
12	475	26.0	1160	3	Sequence 9, Appl
					Sequence 41, Appl

c	13	458.5	25.1	640681	4	US-09-790-988-1	Sequence 1, Appl
	14	402.5	22.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl
	15	402.5	22.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl
	16	377.5	20.7	26289	4	US-09-902-540-1210	Sequence 1210, Ap
	17	339	18.6	546	4	US-09-902-540-4227	Sequence 4227, Ap
	18	318.5	17.4	1959	4	US-09-252-991A-7983	Sequence 7983, Ap
c	19	318.5	17.4	2010	4	US-09-252-991A-8169	Sequence 8169, Ap
c	20	261.5	14.3	601	4	US-09-949-016-50640	Sequence 50640, A
	21	261.5	14.3	109925	4	US-09-949-016-13210	Sequence 13210, A
	22	251.5	13.8	759	4	US-09-902-540-8116	Sequence 8116, Ap
c	23	251.5	13.8	11502	4	US-09-902-540-1064	Sequence 1064, Ap
	24	245	13.4	89892	4	US-09-949-016-13667	Sequence 13667, A
	25	184	10.1	53500	4	US-09-266-965-76	Sequence 76, Appl
c	26	156	8.5	9937	4	US-09-902-540-946	Sequence 946, Appl
	27	155	8.5	6803	3	US-08-665-259-19	Sequence 19, Appl
	28	155	8.5	6803	3	US-08-762-500-19	Sequence 19, Appl
	29	150.5	8.2	483	4	US-09-266-965-69	Sequence 69, Appl
	30	146.5	8.0	2323	3	US-09-149-476-24	Sequence 24, Appl
	31	138.5	7.6	7680	4	US-09-902-540-8948	Sequence 8948, Ap
c	32	138.5	7.6	11476	4	US-09-902-540-955	Sequence 955, Ap
	33	137.5	7.5	1479	4	US-09-252-991A-11768	Sequence 11768, A
	34	137.5	7.5	1968	4	US-09-252-991A-11840	Sequence 11840, A
	35	132.5	7.3	23417	4	US-09-902-540-1207	Sequence 1207, Ap
	36	131.5	7.2	2130	4	US-09-252-991A-5422	Sequence 5422, Ap
	37	127	7.0	58014	4	US-09-949-016-17448	Sequence 17448, A
	38	126.5	6.9	4183	4	US-09-902-540-709	Sequence 709, Appl
c	39	125	6.8	30782	4	US-09-949-016-13724	Sequence 13724, A
c	40	125	6.8	54945	4	US-09-967-669-10	Sequence 10, Appl
	41	124.5	6.8	10106	4	US-09-949-016-17187	Sequence 17187, A
	42	124.5	6.8	27933	4	US-09-949-016-12369	Sequence 12369, A
	43	124	6.8	3446	4	US-09-949-016-1033	Sequence 1033, Ap
	44	124	6.8	3446	4	US-09-949-016-5776	Sequence 5776, Ap
	45	124	6.8	3446	4	US-09-949-016-5777	Sequence 5777, Ap

ALIGNMENTS

RESULT 1

US-09-134-001C-1680

; Sequence 1680, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 1680

; LENGTH: 615

; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-1680

Alignment Scores:					
Pred. No.:	4.17e-36	Length:	615		
Score:	521.50	Matches:	102		
Percent Similarity:	67.33%	Conservative:	34		
Best Local Similarity:	50.50%	Mismatches:	63		
Query Match:	28.56%	Indels:	3		
DB:	3	Gaps:	1		

US-10-829-432-4 (1-343) x US-09-134-001C-1680 (1-615)

Qy 140 lleGlyLysThrAsnIleuTrpHisAnCysLeuIleGlyGlnSerAspArgGln 159

Db 16 ATGAGTGAATCAATCATATTACATGCGATGACTCAGAAAGTTACGAGAAACAAAGACAG 75

```
QY 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 CATAAAAATGGACACAAAAGTCTGTATATGTTTACAGGGCTATCTGGGTGAGGAAA 135

QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 TCACGGGTCTGTGTTGGCTTAGAAAAGAAATTTATTTAATGAAGGAAAAACAACATACCGT 195

QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 TTAGATGGTGTAATGTGGTCTATGCTTAATAAAAAATTTAGGGTTAGTCTCTGAAGAT 255

QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 CGTTCAGAAAATATTCGACGCATTCGAGAAGTAGCAAAATTAATGGTAGACCTGGAGCC 315

QY 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 TTAACAGTTTACTGCTTTTATCTCCCATATAAAGAAGACAGAGAAGGTGTAGAGCATTA 375

QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 CTAGAGGATATGAGTTTATAGAAGTATATACAAAATGTAGTGTGAGGAATGTGAAAAG 435

QY 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 AGAGATCTCTAAAGATTGTATAAAAAGCACCATCTGGGGAAATACCTGAAATTTACAGT 495

QY 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 496 ATAAGTGCACTTATCAACACCTGAAAAACCCCTGAAATTTACTATA-----GATACT 546

QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 GAACAGCACATTTGAACAATCAGTGGTACAAAATTTACGTTTCTTTAAGACACGAA 606

QY 340 TyrLeu 341
    ||| |||
Db 607 TATATT 612
```

RESULT 2

```
US-09-543-681A-2995
; Sequence 2995, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2995
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2995
```

```
Alignment Scores:
Pred. No.: 1.19e-35 Length: 681
Score: 517.00 Matches: 107
Percent Similarity: 60.79% Conservative: 31
Best Local Similarity: 47.14% Mismatches: 57
Query Match: 28.31% Indels: 32
DB: 4 Gaps: 4
```

US-10-829-432-4 (1-343) x US-09-543-681A-2995 (1-681)

```
QY 133 GluLysProValMetSerAsnIleGlyLysSer----- 143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 GATCCGCCGCCATTTTCCCGCATTTGGGGCGCAAGAGACTTACTGGGAGGAAATAGTGAGG 72
```

```
QY 144 -----ThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeu 161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 ATACATCAAGATATGCTGCGCATCTCATCAATAGGTTAAAAAGACGCTGAACACAA 132

QY 162 LeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThr 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 CAGGTACACAAGAGTGTACTTGGTTTACTGGGTATCTGGGTACAGTAAATCAACA 192

QY 182 LeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu----- 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 CTGGTGTATGGCTAGAGCAAACTTATATCAGTACTCGACACTCCATCGCCCTATCCGC 252

QY 197 ThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLys 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 ACCTATTATTAGATGGTGATAATCTAGCCCATGGTTTATGCCCATGATCTTTGGGTTTGT 312

QY 217 AlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAsp 236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 GAACAAGATAGGCATGAAATATTTCGCGTGTAGGGGAAGTGGCTAAATTAATGGTTGAT 372

QY 237 AlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCys 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 GCCGATTAAATTGCTTAAACAGCATTTATTTCTCTTATCAGCAAGATAGACAAAGTA 432

QY 257 ArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIle 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 AGAGAAAGTTTGTCTCAAGGGCGATTATTATTCAGATCTTTGTTGATACACCTTTAGCCCTT 492

QY 277 CysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGly 296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 493 TGTGAAGCACGTGATCTCTAAAGGCTCTATCAAAAAGCGCAGCAGGAGAGATCAAAACAG 552

QY 297 PheThrGlyIleAspAspProTyrGluProPro-----IleAsnGly 310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 TTTTCCGCGCATTGATTCACCTTATGAACCCCCCACTGGCCGAGAAATTCATTTAGACGCG 612

QY 311 GluIleValIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGln 330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 613 TCACCTCGCGATT-----AATGACCTTACACACAA 642

QY 331 ValLeuCysTyrLeuGluGlu 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 643 ATCCTCGCTTATTACAGCAA 663
```

RESULT 3

```
US-09-489-039A-2803
; Sequence 2803, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2803
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2803
```

```
Alignment Scores:
Pred. No.: 1.79e-34 Length: 651
Score: 503.00 Matches: 97
Percent Similarity: 70.24% Conservative: 21
Best Local Similarity: 57.74% Mismatches: 50
Query Match: 27.55% Indels: 0
DB: 4 Gaps: 0
```

US-10-829-432-4 (1-343) x US-09-489-039A-2803 (1-651)

145	Qy	AsnIleLeuTprHisAsnCysLeuIleGlyClnSerAspArgGlnLysIleuLeuGlyGln	164
		: : : : :	
64	Db	AACGTGCTGGCATGCCACCTTGACCCACAGCAGCGGGAACAGCATCACGCCCAT	123
		: : : : :	
165	Qy	LysGlyCysValValTprIleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCys	184
		: : : : :	
124	Db	CGCGCGTGTGTCTGTGGTTTACCGGCGTTTCGGGGTCGGGTAAATCCACCGTCGCCGG	183
		: : : : :	
185	Qy	AlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGlyAspAsn	204
		: : : : :	
184	Db	GGCGTGGAGGAGGCCCTGCATGAGCGCGGTACGACCTATCTCTGACGCGCGATAAC	243
		: : : : :	
205	Qy	LeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIle	224
		: : : : :	
244	Db	GTGCGCCATGGCTGTGTGCAGCGACCTGGGATTTCAGCGATGAGGATCGAAAGAGAACATT	303
		: : : : :	
225	Qy	ArgArgValGlyGluValAlaLysLeuPheAlaAspIleGlyValIleCysIleAlaSer	244
		: : : : :	
304	Db	CGCCCGCTCGCGAAGTGGCCAGGCTGATGTGTCATCCGGGCTGTGGTATTGACGGCA	363
		: : : : :	
245	Qy	LeuIleSerProTyrArgArgAspArgAlaCysArgAlaLeuLeuProHisSerAsn	264
		: : : : :	
364	Db	TTTATATCTCCGACCGTCCGGAACGCAGATGGTGGCGGAGCGCTCGGGGAAGGACGC	423
		: : : : :	
265	Qy	PheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLysGly	284
		: : : : :	
424	Db	TTTATCGAAGTGTTCGTGCATACCCCGTGGCCATCTTCGAGCGCGGGATCCGAAAGA	483
		: : : : :	
285	Qy	LeuTyrLysLeuAlaArgThrGlyLysIleGlyPheThrGlyIleAspAspProTyr	304
		: : : : :	
484	Db	TTGTATAGAAAGCGCGGCGAGGGGAATTACGCAAATTTCACCGGTATAGNCTCGGTCTAT	543
		: : : : :	
305	Qy	GluProIleAsnGlyGluIle	312
544	Db	GAGGCACCGAAAGCGCGAAATT	567

RESULT 4

```

US-09-898-165B-1
; Sequence 1, Application US/09898165B
; Patent No. 6818428
; GENERAL INFORMATION:
; APPLICANT: Daniel H. Cohn
; APPLICANT: Muhammad Faiyaz ul Haque
; APPLICANT: Lily M. King
; APPLICANT: Deborah Krakow
; TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate
; TITLE OF INVENTION: (PAPS) Synthetase Defects and Methods for Treating
; TITLE OF INVENTION: Osteoarthritis Disorders
; FILE REFERENCE: 18810-81552
; CURRENT APPLICATION NUMBER: US/09/898,165B
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/399,212
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-898-165B-1

```

Alignment Scores:	
Pred. No.:	3.81e-33
Score:	496.00
Length:	108
Matches:	108
Percent Similarity:	62.73%
Conservative:	30
Best Local Similarity:	49.09%
Mismatches:	61
Indels:	21
Query Match:	27.16%
DB:	4
Gaps:	5

US-10-829-432-4 (1-343) x US-09-898-165B-1 (1-2014)

114	Qy	ArgArgArgArgArg	-----AsnGlyAlaAlaProGly-GluAlaProHisLe	129
11	Db	CGCGCGCGCGCGGTCCCTCGCTCTTCGGGTCTCTGCTCCCGGGACCCCGGCTCCCGCGCA	70	
129	Qy	rProVallys-----GluLysProValMetSerAsnIleGlyLysSerThrAs	145	
71	Db	CCGACCCAGCATGTCGGGGATCAAGAAGCAAAAGACGGAGAACCCAGCAAAATCCACCA	130	
145	Qy	nIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuGlyGln--	164	
131	Db	TTGTAGTCTATCAGCGCCACCATGTGACGACGAGAAATAAGAGAGGGCAAGTGGTTGGAA	190	
165	Qy	-----LysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThr	181	
191	Db	GGTGGGTTCGAGGAGATGTACCGTGTGGCTAAACAGGTCTCTCTGGTGTCTGGAAACAC	250	
181	Qy	rLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAs	201	
251	Db	GATAAGTTTTGCCCTGGAGGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGA	310	
201	Qy	pGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAl	221	
311	Db	TGGGGCAAAATGTCGCTCATGCGCCTTAACAGAAATCTCGGATTCTCTCTCGGGGACAG	370	
221	Qy	aGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCy	241	
371	Db	GGAAATAATATCCGCGCGATTCCTGAGGTGGCTAAGCTTTGTCTGATGCTGTGTGTG	430	
241	Qy	sIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu----	259	
431	Db	CATTACCAGCTTTATTTCTCCATTTCGCAAAAGGATCGTGAAATGCCCGCAAAATACAT	490	
260	Qy	-----LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysI	276	
491	Db	ATCAGCAGGGCTGCCA-----TTCTTTTGAATAATTGTAGATGCACCTCTAAATAT	541	
276	Qy	eCysGluAlaArgAspProLysGlyLeuTyrlsLysLeuAlaArgThrGlyLysIleLysG	296	
542	Db	TTGTGAAGACGACAGCTAAAGGCCCTCTATAAAGGGCCAGNGCTGGGGAGATTAAAG	601	
296	Qy	yPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValIleLys	315	
602	Db	ATTTACAGGTATTGATCTGATTATGAGAAACCTCGAAATCTCTGACGGTGTGCTTAA	659	

5 T. J. S. A. R.

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```


QY 271 AspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArg 290
Db 475 GATGCTCTCTGCATGTTGTGAACAGAGGATGTCAAGGACTCTACAAAAAGCCGG 534
QY 291 ThrGlyLysIleLysGlyPheThrGlyIleAspAspProTyrGluProProLeuLeuGly 310
Db 535 GCAGGAGAAATTAAGGTTTCACTGGGATCGATTCTGAATATATAAAGCAGAGGCCCT 594
QY 311 GluIleValIleLysMetLysAsp-----GluGluCysProSerProLysAlaMet 327
Db 595 GAGTGTGCTGTAACACACTCTCTGTGATGTAATGACTGT-----636
QY 328 AlaLysGlnValLeuCysTyrLeuGluGlu 337
Db 637 GTCCAGCAAGTTGTGGAAGTTCTTACAGGAA 666

RESULT 7
US-09-949-016-1925
; Sequence 1925, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1925
; LENGTH: 2015
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1925

Alignment Scores:
Pred. No.: 5,13e-33 Length: 2015
Score: 494.50 Matches: 105
Percent Similarity: 64.32% Conservative: 32
Best Local Similarity: 49.30% Mismatches: 59
Query Match: 27.08% Indels: 17
DB: 4 Gaps: 5

US-10-829-432-4 (1-343) x US-09-949-016-1925 (1-2015)

QY 117 ArgArgAsnGlyAlaAlaProGly---GluAlaProHisSerProValLys-----132
Db 31 CGTCTCTCGGTCTCTGCTCCCGGACCCCGGCTCCGCGCAGCCAGCATGTCGGG 90
QY 133 ---GluLysProValMetSerAsnIleGlyLysSerThrAsnIleLeuTyrHisAsnCys 151
Db 91 ATCAGAAGCAAAAGCAGGAGAACAGCAGCAATCCACCANTGTAGTCTATCAGGCCAC 150
QY 152 LeuIleGlyLysSerAspArgGlnLysLeuLeuGlyGln-----LysGlyCys 167
Db 151 CATGTGAGCAGCAATTAAGAGGGCAAGTGTGGAACAAGGGTGGGTTCGAGGATGT 210
QY 168 ValValTyrIleThrGlyLeuSerGlyLysSerThrLeuAlaCysAlaLeuSer 187
Db 211 ACCGTGTGGCTAACAGGTCTCTCTGCTGCTGGGAAACCAACATAGTTTTCCTGGAG 270
QY 188 ArgGluLeuHisCysArgGlyHisLeuThrTyrValIleAspGlyAsnLeuArgHis 207
Db 271 GAGTACCTTGTCTCCATCCATCCCTTGTACTCTCTGCTGGATGGGACATGTCGTCTAT 330
QY 208 GlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgVal 227

Db 331 GGCCTTAACAGAAATCTCGGATTTCTCTCTCGGGACAGAGAGAAATATATCCCGGATT 390
QY 228 GlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSer 247
Db 391 GCTGAGGTGGCTAAAGCTGTTCTGCTGATGCTGCTGGTCTGGCTATACAGCTTTATTCT 450
QY 248 ProTyrArgAspArgAspAlaCysArgAlaLeu-----LeuProHis 262
Db 451 CCATTCGCAAGGATCGTGAGATGCCCGCAAAATACATGAATCAGCAGGCTGCCA---507
QY 263 SerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspPro 282
Db 508 -----TTCTTTGAAATATTTGTAGATGCACCTCTAAATATTTGTGAAGCAGACGTA 561
QY 283 LysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyLeuAsp 302
Db 562 AAAGCCCTCTATAAAGGCGCAGACTGGGGAGATTAAAGGATTATACAGGATTATTCT 621
QY 303 ProTyrGluProProIleAsnGlyGluIleValIleLys 315
Db 622 GATTATGAGAAACCTGAAACTCTCTGAGCGTGTGCTTAAA 660

RESULT 8
US-09-786-240-21
; Sequence 21, Application US/09786240
; Patent No. 6558935
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PF-0592 PCT
; CURRENT APPLICATION NUMBER: US/09/786,240
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 60/133,642
; PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PERL Program
; SEQ ID NO 21
; LENGTH: 2617
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6558935 1420940CB1
US-09-786-240-21

Alignment Scores:
Pred. No.: 3.7e-32 Length: 2617
Score: 486.50 Matches: 104
Percent Similarity: 63.85% Conservative: 32
Best Local Similarity: 48.83% Mismatches: 60
Query Match: 26.64% Indels: 17
DB: 4 Gaps: 5

US-10-829-432-4 (1-343) x US-09-786-240-21 (1-2617)

QY 117 ArgArgAsnGlyAlaAlaProGly---GluAlaProHisSerProValLys-----132
Db 42 CGTCTCTCGGTCTCTGCTCCCGGACCCCGGCTCCGCGCAGCCAGCATGTCGGGG 101
QY 133 ---GluLysProValMetSerAsnIleGlyLysSerThrAsnIleLeuTyrHisAsnCys 151
Db 102 ATCAGAAGCAAAAGCAGGAGAACAGCAGCAATCCACCANTGTAGTCTATCAGGCCAC 161
QY 152 LeuIleGlyLysSerAspArgGlnLysLeuLeuGlyGln-----LysGlyCys 167

Db 162 CATGTGAGCAGGATAAGAGAGGGCAAGTGGTTGGAAACAAGGGGTGGGTTCGAGGATGT 221
Qy 168 ValValTrpIleThrGlyLeuSerGlySerGlyLeuSerThrLeuAlaCysAlaLeuSer 187
Db 222 ACCGTGTGGCTAACAGGTCTCTCTGGTGTGGAACACCAACATAGATTTTGGCCCTGGAG 281
Qy 188 ArgGluLeuHisCysArgGlyHisLeuThrTrpValLeuAspGlyAspAsnLeuArgHis 207
Db 282 GAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGGCAATGTCCGTCAAT 341
Qy 208 GlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgVal 227
Db 342 GCGCTTAACAGAAATCTCGATCCCTCTCTGGGGACAGAGAGAAAATATCCGCCGAT 401
Qy 228 GlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSer 247
Db 402 GCTGAGGTGGCTAAGCTGTCTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 461
Qy 248 ProTyrArgArgAspAlaCysArgAlaLeu-----LeuProHis 262
Db 462 CCATTCGCAAGGATCGTGAGATGCCCGCAAAATACATGATCAGCAGGGCTGCCA--- 518
Qy 263 SerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspPro 282
Db 519 -----TTCCTTTGAAATATTTGTAGATGCACCTCTAAATATTTGTGAAAGCAGACGTA 572
Qy 283 LysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAsp 302
Db 573 AAAGGCTCTATAAAGGCCAGAGCTGGGGAGATTAAGGATTTACAGGATTAATGATCT 632
Qy 303 ProTyrGluProPheIleAsnGlyGluIleValIleLys 315
Db 633 GATTATGAGAAACCTTGAAACTCTGAGCGTGTGCTTAAA 671

RESULT 9

US-09-898-165B-10
; Sequence 10, Application US/09898165B
; Patent No. 6818428
; GENERAL INFORMATION:
; APPLICANT: Daniel H. Cohn
; APPLICANT: Muhammad Faiyaz ul Haque
; APPLICANT: Lily M. King
; APPLICANT: Deborah Krakow
; TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate
; TITLE OF INVENTION: (PAPS) Synthetase Proteins and Methods for Treating
; TITLE OF INVENTION: Osteoarthritis Disorders
; FILE REFERENCE: 18810-81552
; CURRENT APPLICATION NUMBER: US/09/898,165B
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/399,212
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1851
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-898-165B-10

Alignment Scores:
Pred. No.: 3,31e-32 Length: 1851
Score: 484.50 Matches: 101
Percent Similarity: 65.37% Conservative: 33
Best Local Similarity: 49.27% Mismatches: 56
Query Match: 26.53% Indels: 15
DB: 4 Gaps: 4

US-10-829-432-4 (1-343) x US-09-898-165B-10 (1-1851)

Qy 142 LysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeu 161
Db 43 AAATCCACCAATGTGGTCTACAGGCCCTCATGTGAGCAGGAAACAAGAGACGAATGTG 102

Qy 162 LeuGlyGln-----LysGlyCysValValTrpIleThrGlyLeuSerGlySer 177
Db 103 GTTGGAACACGGGAGGATTCGAGGATGTACCGTGTGGCTAACAGGTCTCTCTGGTGTCT 162
Qy 178 GlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThr 197
Db 163 GGGAAAACAACCAATAGCTTTGGTGGAGAGTACCTTTGATCTCATCGCCATCCCATGT 222
Qy 198 TyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAla 217
Db 223 TACTCCCTGGATGGGACCAATGTCCGTCAATGSCCTTAATAAGAACCTGGGATTTCTTCC 282
Qy 218 GluAspArgAlaGluAsnIleArgValGlyGluValAlaLysLeuPheAlaAspAla 237
Db 283 GGGGACCCGAGAAGAGATATATCCGCGGATCCGCGAGGTGGCCAAAGCTCTTTGCCGACGCC 342
Qy 238 GlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAlaCysArg 257
Db 343 GGCCTGGTTGTCATCACCAGCTTTATCTCTCTTTTGCAAAAGGATCGTGAGAAATGCCGA 402
Qy 258 AlaLeu-----LeuProHisSerAsnPheIleGluValPheIleAspLeu 272
Db 403 AAATCCACGAAATCAGCAGGACTCCG-----TTCCTTGGATCTTTGTAGATCGG 453
Qy 273 ProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGly 292
Db 454 CTTTAAATATCTGTGAAAGCCGAGACGTAAAGGACTCTACAAACGAGCCGACGAGCA 513
Qy 293 LysIleLysGlyPheThrGlyIleAspAspProTyrGluProPheIleAsnGlyGluIle 312
Db 514 GAGATTAAGGGTTTACAGGATCGATTTCTGACTATGAGAAACCTTGAAATCCAGAGTGT 573
Qy 313 ValIleLysMetLysAspGluCysProSerProLysAlaMetAlaLysGlnValLeu 332
Db 574 GTGCTGAG-----ACCAACTGTCTTCAGTACGCGACTGTGTCCACAGGTGTGTG 624
Qy 333 CysTyrLeuGluGlu 337
Db 625 GAACTTTTGCAGGAG 639

RESULT 10

US-09-898-165B-2
; Sequence 2, Application US/09898165B
; Patent No. 6818428
; GENERAL INFORMATION:
; APPLICANT: Daniel H. Cohn
; APPLICANT: Muhammad Faiyaz ul Haque
; APPLICANT: Lily M. King
; APPLICANT: Deborah Krakow
; TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate
; TITLE OF INVENTION: (PAPS) Synthetase Proteins and Methods for Treating
; TITLE OF INVENTION: Osteoarthritis Disorders
; FILE REFERENCE: 18810-81552
; CURRENT APPLICATION NUMBER: US/09/898,165B
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/399,212
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-898-165B-2

Alignment Scores:
Pred. No.: 3,71e-32 Length: 2000
Score: 484.50 Matches: 101
Percent Similarity: 65.37% Conservative: 33
Best Local Similarity: 49.27% Mismatches: 56
Query Match: 26.53% Indels: 15
DB: 4 Gaps: 4

US-10-829-432-4 (1-343) x US-09-898-1658-2 (1-2000)

QY 142 LysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyClnSerAspArgGlnLysLeu 161
 Db 103 AATCCACCATGTGCTTACAGGCCCATCATGTGAGCAGGAACAAAGAGAGACAAAGTG 162
 QY 162 LeuGlyGln-----LysGlyCysValValTrpIleThrGlyLeuSerGlySer 177
 Db 163 GTTGGAAACACGAGGAGATTCCGAGGATGTACCGTGTGCTTAACAGGTCTCTCTGTGTCT 222
 QY 178 GlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThr 197
 Db 223 GGGAAACACCAATAGCTTTGCTTGGAGAGTACCTTGTATCTCACGCCATCCCATGT 282
 QY 198 TyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAla 217
 Db 283 TACTCCCTGTGATGGGACCAATGTCCGTCAATGGCTTAAAGAACCTGGGATTTCTGTGCC 342
 QY 218 GluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAla 237
 Db 343 GGGGACCGAGAGAGAAATATCCGCGGATCGCGGAGGTGGCCAAAGCTCTTTGCCGAGCC 402
 QY 238 GlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArg 257
 Db 403 GGCCTGGTTGTCATCACAGCTTTATCTCTCTCTTTTGAAGGATCGTGAGAAATGCCCGA 462
 QY 258 AlaLeu-----LeuProHisSerAsnPheIleGluValPheIleAspLeu 272
 Db 463 AATATCCAGATATCACAGGACTCCCG-----TTCTTGAGATCTTTGAGATGGC 513
 QY 273 ProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGly 292
 Db 514 CCTTTAAATATCTGTGAAGCGAGACGTAAGAAAGACTCTACAAACGAGCCGAGCAGGA 573
 QY 293 LysIleLysGlyPheThrGlyLeuAspAspProTyrGluProPheLeuAsnGlyGluIle 312
 Db 574 GAGATTAAAGGGTTTACAGGCATCGATTTCTGACTATGAGAAACCTGAAACTCCAGAGTGT 633
 QY 313 ValIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeu 332
 Db 634 GTGCTGAAG-----ACCAACTGTCTTCAGTAAGCGACTGTGTGAACAGGTGGTG 684
 QY 333 CysTyrLeuGluGlu 337
 Db 685 GAATTTTTCAGGAG 699

RESULT 11

US-09-898-1658-9
 ; Sequence 9, Application US/098981658
 ; Patent No. 6818428
 ; GENERAL INFORMATION:
 ; APPLICANT: Daniel H. Cohn
 ; APPLICANT: Muhammad Faiyaz ul Haque
 ; APPLICANT: Lily M. King
 ; APPLICANT: Deborah Krakow
 ; TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate
 ; TITLE OF INVENTION: (PAPS) Synthetase Proteins and Methods for Treating
 ; TITLE OF INVENTION: Osteoarthritic Disorders
 ; FILE REFERENCE: 18810-81552
 ; CURRENT APPLICATION NUMBER: US/09/898,1658
 ; CURRENT FILING DATE: 2001-07-02
 ; PRIOR APPLICATION NUMBER: 09/399,212
 ; PRIOR FILING DATE: 1999-09-17
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 1845
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-898-1658-9

Alignment Scores:

Pred. No.: 5.42e-32 Length: 1845

Score: 482.00 Matches: 98
 Percent Similarity: 66.67% Conservative: 30
 Best Local Similarity: 51.04% Mismatches: 52
 Query Match: 26.40% Indels: 12
 DB: 4 Gaps: 3

US-10-829-432-4 (1-343) x US-09-898-1658-9 (1-1845)

QY 133 GluLysProValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeu 152
 Db 13 AAGAAGCAAAAGACCGGAGAACCCAGAGAAATCCCAATGTAGTCTATCAGGCCCAACCAT 72
 QY 153 IleGlyGlnSerAspArgGlnLysLeuGlyGln-----LysGlyCysVal 168
 Db 73 GTGACGAGGAATAAGAGAGGGCAAGTGTGGCAACAGGGGTGGTTCGAGGATGTACC 132
 QY 169 ValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArg 188
 Db 133 GTGTGGCTAACAGGCTCTCTGTGTCTGGAACCAACGATTAAGTTTTCCTCGGAGGAG 192
 QY 189 GluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGlyAspAsnLeuArgHisGly 208
 Db 193 TACCTGTCTCCCATCCCATCCCTGTCTCTCCCTGGATGGGACAAATGTCCGTATGCG 252
 QY 209 LeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgArgValGly 228
 Db 253 CTTAACAGAAATCTCGGATTTCTCTCTGGGACAGAGAGAAATATCCGCCGATGTCT 312
 QY 229 GluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerPro 248
 Db 313 GAGGTGGCTAAGCTCTTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 372
 QY 249 TyrArgArgAspArgAspAlaCysArgAlaLeu-----LeuProHisSer 263
 Db 373 TTCCGAAAGGATCGTGAGAAATGCCCGCAAAATACATGATATCATGAGGGCTGCCA 426
 QY 264 AsnPheIleGluValPheIleAspLeuProLysIleCysGluAlaArgAspProLys 283
 Db 427 ---TTCTTTGAAATATTTGTAGATGCACCTCTAAATATTTGTGAAGCAGAGACGTA 483
 QY 284 GlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyLysAspAspPro 303
 Db 484 GGCCTCTATAAAGGCGCAGAGCTGGGAGATTAAGGATTTACAGGTATTGATCTGTAT 543
 QY 304 TyrGluProPheLeuAsnGlyGluIleValIleLys 315
 Db 544 TATGAGAAACCTGAAACTCCTGAGCGGTGTGTCTTAAA 579

RESULT 12

US-09-153-310-41
 ; Sequence 41, Application US/09153310
 ; Patent No. 6326184
 ; GENERAL INFORMATION:
 ; APPLICANT: Gjermansen, Claes
 ; Hansen, Jorgen
 ; Johannesen, Pia Francke
 ; Pedersen, Mogens Bohl
 ; Sorensen, Steen Bech
 ; TITLE OF INVENTION: Method of producing a composite
 ; fermented beverage using genetically modified yeast
 ; strains
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W.
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS

```
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/153,310
/ FILING DATE: 15-Sep-1998
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: <Unknown>
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bent, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: <Unknown>
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-672-5300
/ TELEFAX: 202-672-5399
/ TELEX: <Unknown>
/ INFORMATION FOR SEQ ID NO: 41:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1160 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: Genomic DNA
/ SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-153-310-41
```

```
Alignment Scores:
Pred. No.: 1,1e-31 Length: 1160
Score: 475.00 Matches: 99
Percent Similarity: 66.8% Conservative: 24
Best Local Similarity: 53.8% Mismatches: 55
Query Match: 26.0% Indels: 6
DB: 3 Gaps: 3
```

US-10-829-432-4 (1-343) x US-09-153-310-41 (1-1160)

```
QY 143 SerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeu 162
   ::::::::::::::::::::
Db 381 GCTACTAATATCACTGGCATCCAAATCTTACC---TACGACGACGATGAAGAAATGA 437

QY 163 GlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeu 182
   ::::::::::::::::::::
Db 438 AAGCAAGACGGCTGTACCGCTTGGTTCACCGCTCAAGTGCCTCAGGAAAAAGTACAATA 497

QY 183 AlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGly 202
   ::::::::::::::::::::
Db 498 GCTTGTGCACTGGACAACTACTCTCTCAAAAAAACTTATCTGCTTATAGGTAGATGT 557

QY 203 AspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGlu 222
   ::::::::::::::::::::
Db 558 GATAACATTCGTTTGGTTTGAATAAGGATTTGGGCTTCTCAGAAAAAGGACAGAAATGAA 617

QY 223 AsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIle 242
   ::::::::::::::::::::
Db 618 AACATTCGTAGAAATAGTAGAAGTATCCAAAGCTATTGCGTGAATTCGTGCTGTATCCATC 677

QY 243 AlaSerLeuIleSerProTyrArgArgAspAlaCysArgAlaLeuLeuProHis 262
   ::::::::::::::::::::
Db 678 ACTTCATTTATTTCCCATACAGATCGATAGACAGACAGCCGCGTGAATTTACATAAGNA 737

QY 263 Ser-----AsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArg 280
   ::::::::::::::::::::
Db 738 GCAGGGCTTCAAGTTCAATGAAATTTTGTGTATGTTCCATTAGAAGTCGCTGAGCAAGA 797

QY 281 AspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIle 300
   ::::::::::::::::::::
Db 798 GACCCCTAAGGGTTGTATAGAAAGCCAGAGAAGGTGTGATTAAAGAGTTCTACTGGTATT 857

QY 301 AspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAsp----- 318
   ::::::::::::::::::::
Db 858 TCAGTCTCCTTACGAAGCTCCAAAGGCCCGCAGAGTTGCTATTAAAGAACTGACCAAGACT 917

QY 319 ----GluGluCys 321
```

```
Db 918 GTTGAAGAATCT 929
RESULT 13
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1
```

```
Alignment Scores:
Pred. No.: 3,19e-26 Length: 640681
Score: 458.50 Matches: 99
Percent Similarity: 60.77% Conservative: 28
Best Local Similarity: 47.37% Mismatches: 75
Query Match: 25.11% Indels: 7
DB: 4 Gaps: 2
```

US-10-829-432-4 (1-343) x US-09-790-988-1 (1-640681)

```
QY 127 ProHisSerProValLysGluLysProValMetSerAsnIleGlyLys----- 142
   ::::::::::::::::::::
Db 457846 CCACATTTGGAACATA---CCAAAAAATTTTAATGAAAAAGGTTTATAAGTAATAAACAAT 457790

QY 143 -----SerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLys 160
   ::::::::::::::::::::
Db 457789 AATTTTCAAAAACAATATTTTTCGCAAAAACATTCATTCACGCTTAAAAACGTAAGAAA 457730

QY 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer 180
   ::::::::::::::::::::
Db 457729 AAAAATGCTCAATAATCAATTTGTTACTATGTTCTACTGGCTCTCAGGTTTCAGAAAAATCA 457670

QY 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeu 200
   ::::::::::::::::::::
Db 457669 ACCATAGCAAAATTTTGTAGAAGAAATATTTTAAAAAATGGAATTAATAGCTATTATTA 457610

QY 201 AspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArg 220
   ::::::::::::::::::::
Db 457609 GACGAGACAATATAGGTCTGGTTTATGTTCTGATTTAAGTTTTTAGTTTGGCAGATAGG 457550

QY 221 AlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIle 240
   ::::::::::::::::::::
Db 457549 AATGAAAACATTTAGACGATTTGGAGAACTAGTAAAAATGATGTTACATGCTGTTCTAATA 457490

QY 241 CysIleAlaSerLeuIleSerProTyrArgArgAspAlaCysArgAlaLeuLeu 260
   ::::::::::::::::::::
Db 457489 ATATTGGTATCAGTTATTTCCCTCTATAGAAATCAAAAGGGAATGGTACGTCAATATGTTA 457430

QY 261 ProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArg 280
   ::::::::::::::::::::
Db 457429 GGAAAAAAACTTTTATAGAAAGTTTTCATGATACACCTATTGAATTTTGAAATCTCGA 457370

QY 281 AspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIle 300
   ::::::::::::::::::::
Db 457369 GATCCTAAAAAATATATAAAACAGCCGACAGGCCCAAAATATCTGATTTACTGTTATA 457310

QY 301 AspAspProTyrGluProIleAsnGlyGluIleValIleLysMetLysAspGluGlu 320
```



```
Db 457309 CAATGTACATACGAAACTCCTATACACCTGATGTTCTTTAAAGGGTACAGATTCTTTA 457250
QY 321 CysProSerProLysAlaMetAlaLys 329
Db 457249 AAAAATAACTCAAAAAAATTATATAAA 457223

RESULT 14
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 3,74e-20 Length: 4403765
Score: 402.50 Matches: 120
Percent Similarity: 44.38% Conservative: 34
Best Local Similarity: 34.58% Mismatches: 111
Query Match: 22.04% Indels: 83
DB: 3 Gaps: 11

US-10-829-432-4 (1-343) x US-09-103-840A-2 (1-4403765)
QY 11 ProLeuValThrHisThrGlnGlnProSerProAlaProGlyProAlaSerGlnGly 30
Db 1439362 CCGGCGCGACTACGTTGTCAAGCACCA-----CCGAACCGTCCGCGGAGGA 1439412
QY 31 GlnArgGlnGlyAsnThrLeuLeuSerProThrProThrLeuAlaValileLeuValAsn 50
Db 1439413 TAGCCGGGCTGGATTACCGGCTCGATGTCAACACCGCTGCATCGCG-----ACAAAGA 1439463
QY 51 ProGlnArg-----AlaProProValLeu 58
Db 1439464 CCGCAACGGCGTTGAACACGAGTGGCGGTTCGTCGCGACCCAGGTCGCGT 1439523
QY 59 ProGlyLeuThrProSerAspAlaProLeuProAlaLeuValileHisGlyLeuThrPro 78
Db 1439524 TCGTCTTACGAGTACACCGGACCGTAGCACCGGCTGCTTCATCTCATGACCCCG 1439583
QY 79 ArgSerSerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgGlyGlyGly 98
Db 1439584 ACACCAACGGA-----1439595
QY 99 ArgGlyAlaArgThrHisCysHisArgGlyIleGlyArgTTPValArgArgArg 118
Db 1439596 -----CGGTGG-----CGGCGG 1439607
QY 119 AsnGlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSer 138
Db 1439608 GCATGTGTACCGGACGCTCGGCCCGC-AGCCT-----1439642
QY 139 AsnIleGlyLeuSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyLeuSerAspArg 158
Db 1439643 -----AGCCCGAACACACGCGTGGCGCACAGATCGCTCGTCACTGCGCAAGATCGG 1439690
QY 159 GlnLysLeuLeuGlyGlnLysGlyCysValValTTPileThrGlyLeuSerGlySerGly 178
Db 1439691 CCG-----CCGAGGGGCAAGACGGTGTGTTTACCGGACTGTTCGGGCTCCGGC 1439738
QY 179 LysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyr 198
Db 1439739 AAGTCGTGGTGGCCATGCTGTTGAGCGGAAGCTACTCGAAAAGGCGATCTCCGCTTAC 1439798
QY 199 ValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelysAlaGlu 218
Db 1439799 GTTCTGAGCGGCGACCACTACGGCATGCGCTCAACGCGCGCTTTTCCATGGCC 1439858
QY 219 AspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGly 238
Db 1439859 GACCGCGGGAGAACCTGCGCGGCTGCGCATGTGGGCACACTGCTGCCCATTTGTGGC 1439918
QY 239 ValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAlaCysArgAla 258
Db 1439919 CACCTGGTGTGGTGGCGCGCATCAGCCCTTGTGTGAGCACCGTGGCTCGTAA 1439978
QY 259 LeuLeuProHisSer-----AsnPhelleGluValPheIleAspLeuProLeu 274
Db 1439979 GTG-----CACGCTGATCGCGGAATCGACTTTTTCGAGGTGTCTGTGACACCCCGCTG 1440032
QY 275 LysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIle 294
Db 1440033 CAGGACTGTGAGCGGTGATCCCAAGGGTTGTAGCCCAAGCCGCGTGGGTGAGATC 1440092
QY 295 LysGlyPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValIle 314
Db 1440093 ACGCACTTCACCGGGATCGACAGCCCATATCAGCGGCCCAAGAACCCAGACCTACGGCTT 1440152
QY 315 -----LysMetLysAspGluCysProSerProLysAlaMetAlaLysGlnVal 331
Db 1440153 ACGCGGATCGCAGCATAGACGAGCAG-----GCCGAGGAGGTT 1440191
QY 332 LeuCysTyrLeuGluGluAsn 338
Db 1440192 ATCGACCTGTTGGAGTCAATCG 1440212

RESULT 15
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 3,75e-20 Length: 4411529
Score: 402.50 Matches: 120
Percent Similarity: 44.38% Conservative: 34
Best Local Similarity: 34.58% Mismatches: 111
Query Match: 22.04% Indels: 83
DB: 3 Gaps: 11
```

US-10-829-432-4 (1-343) x US-09-103-840A-1 (1-4411529)

Search completed: September 9, 2005, 15:57:40
Job time : 2514 secs

```
QY      11  ProLeuValThrHisThrGlnGlnProProSerProAlaProGlyProAlaSerGlnGly 30
Db      1439893  CCGGCCGCGACTACCTGTGTAAGCACACCA-----CCGAACCGTCCGCGCGAGGA 1439943
QY      31  GlnArgGlnGlyAsnThrLeuLeuSerProThrProThrLeuAlaValIleLeuValAsn 50
Db      1439944  TAGCCGGGCTGATTACCGGCTCGATGTCAACACCCTGCATCGCG-----ACAAGA 1439994
QY      51  ProGlnArg-----AlaProProValLeu 58
Db      1439995  CCGCAACGGCGGTTGAAACTCAACGAACCTGGGCGGTGTTTCGCTGGCGCACCCAGGTGCCGT 1440054
QY      59  ProGlyLeuThrProSerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrPro 78
Db      1440055  TCCTGCTTGACGAGTACACCCCGCAACCGGTAGCACCGGCTCGTTCATCTCATTTGACCCCG 1440114
QY      79  ArgSerSerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgGluGlyGluGly 98
Db      1440115  ACACCAACGGAA----- 1440126
QY      99  ArgGlyAlaArgThrHisCysHisArgGlyIleGlyArgTTPValArgArgArgArg 118
Db      1440127  -----CGGTGG-----CGGCGG 1440138
QY     119  AsnGlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSer 138
Db      1440139  GCATGGTGTTACGCGACGCTCGGCCCGC-ACGCCT----- 1440173
QY     139  AsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArg 158
Db      1440174  -----AGCCCGACACGGTGGCGGCACAGATCGCTCGTCACTCGCGCAAGATCGG 1440221
QY     159  GlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGly 178
Db      1440222  CCG-----CCACGGGGCAAGACGGTGTGTGTTCACCGGACTGTCCGGCTCCGCGC 1440269
QY     179  LysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyr 198
Db      1440270  AAGTCGTGGTGCCATGCTGTTGAGCGGAAGTACTCGAAAAGGGCATCTCCGCTTAC 1440329
QY     199  ValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelysAlaGlu 218
Db      1440330  GTTCTGGACGGCGACAACTACGGCATGGCCTCAACGCCGACCTGGGCTTTTCCATGGCC 1440389
QY     219  AspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGly 238
Db      1440390  GACCGCGCGAGAACTCGCGCGGCTGTGCGCATGTGGCCACACTGTCTCGCGATTGTGGC 1440449
QY     239  ValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAlaCysArgAla 258
Db      1440450  CACCTGGTCTGTGGTCCCGCATACGCCCTTGTCTGAGCACCGCTGGCTCGTAA 1440509
QY     259  LeuLeuProHisSer-----AsnPheIleGluValPheIleAspLeuProLeu 274
Db      1440510  GTG-----CACGCTGATCGGGGATCGACTTTTTCGAGGTGTCTGTGACACCCCGCTG 1440563
QY     275  LysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIle 294
Db      1440564  CAGGACTGTGAGAGGCGTGATCCCAAGGGTTGTACGCCAAAGCGCGTGGGTGAGATC 1440623
QY     295  LysGlyPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValIle 314
Db      1440624  ACGCACTTACCGGGATCGACAGCCCATATACGGGCCCAAGAACCCAGACCTACGGCTT 1440683
QY     315  -----LysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnVal 331
Db      1440684  ACGCCGGATCGCAGCATACGAGCAG-----GCCGAGGAGGTT 1440722
QY     332  LeuCysTyrLeuGluAsn 338
Db      1440723  ATCGACCTGTGTGAGTCATCG 1440743
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2005, 12:40:30 ; Search time 47 Seconds
(without alignments)
702.178 Million cell updates/sec

Title: US-10-829-432-4
Perfect score: 1826
Sequence: 1 RPFHFINQTEPLVTHTQPP.....PRMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	833	45.6	312	2 T08076	adenyl-yl-sulfate k
2	816	44.7	276	1 S47640	adenyl-yl-sulfate k
3	777.5	42.6	293	2 T08100	adenyl-yl-sulfate k
4	571	31.3	200	2 E96912	adenyl-yl-sulfate ki
5	555	30.4	199	2 A69839	adenyl-yl-sulfate ki
6	534.5	29.3	202	2 A83836	adenyl-yl-sulfate ki
7	511.5	28.0	633	2 H95932	probable adenyl-yl-
8	510	27.9	213	2 AF0408	adenyl-yl-sulfate k
9	508.5	27.8	635	2 A87433	hypothetical prote
10	505	27.7	652	1 T24918	3'-phosphoadenosin
11	503.5	27.6	202	2 T50101	adenyl-yl-sulfate ki
12	502.5	27.5	610	1 JCA383	adenyl-yl-sulfate k
13	498.5	27.3	660	2 G82672	ATP sulfurylase, l
14	495	27.1	201	2 AH0856	adenosine 5'-phosp
15	495	27.1	624	1 JW0087	adenyl-yl-sulfate k
16	494.5	27.1	215	2 F82062	adenyl-yl-sulfate k
17	492	26.9	644	2 AD3471	adenyl-yl-sulfate k
18	490.5	26.9	641	1 Z22RNO	adenyl-yl-sulfate k
19	490.5	26.9	641	2 E95320	adenyl-yl-sulfate k
20	481	26.3	201	1 B65056	adenyl-yl-sulfate k
21	481	26.3	201	2 D91079	adenosine 5'-phosp
22	481	26.3	201	2 E85924	adenosine 5'-phosp
23	481	26.3	197	2 A84073	adenyl-yl-sulfate ki
24	480	26.3	198	2 C69877	adenyl-yl-sulfate k
25	478.5	26.2	620	1 I39755	adenyl-yl-sulfate k
26	474	26.0	202	2 H17244	adenyl-yl-sulfate k
27	457	25.0	206	2 H84978	adenyl-yl-sulfate k
28	448.5	24.6	196	2 H83472	adenosine 5'-phosp
29	402	22.0	614	1 B70772	probable adenyl-yl-

30	398	21.8	546	1 C70393	probable adenyl-yl-
31	387	20.2	177	1 S74917	adenyl-yl-sulfate k
32	369	20.2	214	1 S18729	adenyl-yl-sulfate k
33	351	19.2	192	2 B75594	adenyl-yl-sulfate ki
34	325	17.8	186	2 G72590	probable adenyl-yl-
35	322.5	17.7	574	1 S55034	sulfate adenyl-yl-
36	314.5	17.2	174	2 F75097	adenyl-yl-sulfate 3-
37	307	16.8	573	1 A53651	sulfate adenyl-yl-
38	305	16.7	155	2 H69285	adenyl-yl-sulfate 3-
39	302	16.5	633	2 D83091	ATP sulfurylase GT
40	179	9.8	170	2 G81286	probable adenyl-yl-
41	111.5	6.1	396	2 T36678	probable septum ei
42	108	5.9	1016	2 D86308	translation initia
43	106	5.8	478	2 T10030	hypothetical prote
44	106	5.8	586	2 H86914	conserved hypothet
45	106	5.8	723	2 B38749	3-phosphatidylnos

ALIGNMENTS

RESULT 1

T08076
adenyl-yl-sulfate kinase (EC 2.7.1.25) precursor - Madagascar periwinkle
N:Alternate names: adenosine-5'-phosphosulfate-kinase
C:Species: Catharanthus roseus (Madagascar periwinkle)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C:Accession: T08076
R:Arz, H.E.; Gisselmann, G.; Schiffmann, S.; Schwenn, J.D.
Biochim. Biophys. Acta 1218, 447-452, 1994
A:Title: A cDNA for adenyl-yl sulphate (APS)-kinase from Arabidopsis thaliana.
A:Reference number: S47640; MUID:94325358; PMID:8049272
A:Accession: T08076
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-312 <ARZ>
A:Cross-references: UNIPROT:O49204; EMBL:AF044285; NID:g28322299; PIDN:AAC31145.1; PID:g2
C:Genetics:
A:Gene: akn
C:Function:
A:Description: phosphorylates 3'-OH group of adenosine-5'-phosphosulfate
C:Superfamily: adenyl-yl-sulfate kinase; adenyl-yl-sulfate kinase homology
C:Keywords: ATP; P-loop; phosphotransferase; purine nucleotide binding
F:135-298/Domain: adenyl-yl-sulfate kinase homology <ASK>

Query Match	45.6%	Score	833	DB	2	Length	312
Best Local Similarity	72.9%	Fred. No.	2.2e-55				
Matches	156	Conservative	24	Mismatches	34	Indels	0
Gaps	0						
Qy	130	PVKKEPVMNSNIGKSTNIIWHNCLIGQSDRQKLLGQKGVWVITGLSGSGKSTLACALSRG	189				
Db	99	PGKILQTTTGVNSTNIIWHKCAVEKSERQEPLOQRGCVIWTGLSGSGKSTLACALSRG	158				
Qy	190	LHCRGHLYVLDDGNLRHGLNRDLSFKAEDRAENIRRVGEVAKUFADAGVTCIASLISPY	249				
Db	159	LHAKGKLYTILDDGNVRHGLNSDLSFKAEDRAENIRRVGEVAKUFADAGVTCIASLISPY	218				
Qy	250	RRDRDACCALLPHSNFIEVFIDLPLKICEARDPKLYKARTGKTKGTGDDDPYEPPIP	309				
Db	219	RKPPDACRSLPEGDFIEVFMDVPLKVCCEARDPKLYKARAGKIKGTGIDDPYEPPLK	278				
Qy	310	GEIVKMKDECPSPKAMAKOVLCVLENGYLQA	343				
Db	279	SEIVLHQLKGMCDSPCLADIVISYLENGYLKA	312				

RESULT 2

S47640
adenyl-yl-sulfate kinase (EC 2.7.1.25) precursor - Arabidopsis thaliana
N:Alternate names: APS kinase; protein F26C24.11
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S47640; T02601; A84521; S38587

C.Species: Bacillus halodurans
C.Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C.Accession: A83836
R.Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Fujii, F.; Hirai,
Nucleic Acids Res. 28, 4317-4331, 2000
A.Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A.Reference number: A83650; MUID:20512582; PMID:11058132
A.Accession: A83836
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-202 <STO>
A.Cross-references: UNIPROT:Q9KCT0; GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BAB052
A.Experimental source: strain C-125
C.Genetics:
A.Gene: BH1489
C.Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology

	Query Match	29.3%	Score 534.5;	DB 2;	Length 202;
	Best Local Similarity	53.8%;	Pred. No. 3.9e-33;	Mismatches 106;	Conservative 28;
					Gaps 2;
QY	145	NILWENCILIGSDROKLLGQKGVVITGLSGSGKSLACALSRELHCRGHLYTVLDGDN	204		
Db	8	HIVHEASVSKSEERQRNLKSCVVWFVLGSFGKSTLANADRLKFEEIGHSYVLDGN	67		
QY	205	LRHLNRDLSPKAEDRAENIRRVGEVAKLFDAGVICIASLISPYRRDRDACLPHSN	264		
Db	68	IRHGLNAGLGFSEEDKENIRRIEAVAKLFVDAGVVTAFISPFREDRDNVRGILDDE	127		
QY	265	FIEVFIDLPLKICEARDPKGLYKLARTGKIKGFTGIDDPEPPINGEIVIKMKDECPSP	324		
Db	128	FIEVYVRCPLEFCERDKPGLYKKARSGDIPEFTGISSPYESPVNPETII---DTDQLAV	184		
QY	325	KAMAKOVLCYL--EENG 339			
Db	185	EAEVEKIYAYLHAQESG 201			

RESULT 7
H95932
Probable adenylyl-sulfate kinase (EC 2.7.1.25) [imported] - Sinorhizobium meliloti (stra
C.Species: Sinorhizobium meliloti
C.Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C.Accession: H95932
R.Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A.Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A.Reference number: A95842; MUID:21396508; PMID:11481431
A.Accession: H95932
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-633 <KUR>
A.Cross-references: UNIPROT:Q92VH6; GB:AL591985; PIDN:CAC49128.1; PID:g15140613; GSPDB:(
A.Experimental source: strain 1021, megaplasmid pSymB
R.Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.,
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A.Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
Hebaut, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A.Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A.Reference number: A96039; MUID:21368234; PMID:11474104
A.Contacts: annotation
C.Genetics:
A.Gene: nodQ2; SMB21224
A.Genome: plasmid
C.Superfamily: modulation protein nodQ; adenylylsulfate kinase homology; translation elc
C.Keywords: phosphotransferase

Query Match 28.0%; Score 511.5; DB 2; Length 633;
Best Local Similarity 54.0%; Pred. No. 8.3e-31;
Matches 101; Conservative 24; Mismatches 61; Indels 1

[illegible]

RESULT 13
G82672
ATP sulfurylase, large subunit XF1501 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C;Accession: G82672
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: G82672
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-660 <STM>
A;Cross-references: GB:AE003980; GB:AE003849; NID:g9106531; PIDN:AAF84310.1; GSPDB:GN00
R;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer,
as-Neto, E.; Docena, C.; El-Dorriy, H.; Faccinani, A.P.; Ferreira, A.J.S.

[illegible]

```
Db 549 RYAEVARLWVDAAGLVVLSFISPFDRQLARERFAANEFEVFDVPLAVAEARDVKGL 608
QY 286 VKLARTGKIKGTGIDDDPYEPINGEIVIKMKDECPSPKAMAKQVLCYL 335
Db 609 YAKARAGLITDTGIDSPYEPHPHDLUR---ADQGTPEQLASQVLSFL 655

RESULT 14
AH0856
adenosine 5-phosphosulfate kinase [imported] - Salmonella enterica subsp. enterica serov
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AH0856
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AH0856
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-201 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD06039.1; PID:gl6504006; GSPDB:GN00176
C:Genetics:
A:Gene: cysC
C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology

Query Match 27.1%; Score 495; DB 2; Length 201;
Best Local Similarity 54.0%; Pred. No. 3.6e-30;
Matches 95; Conservative 29; Mismatches 46; Indels 6; Gaps 1;

QY 145 NILHNCILIGSDRQKLLGQGVVITGLSGSGKSTLACALSRELHCRGHLYTVLVDGN 204
Db 7 NVVHSHPTVVAAREQLGHRGVLVWFTGLSGSGKSTVAGALEALHQGVSTYLLDGN 66

QY 205 LRHGLNRDLSPKAEDRAENIRVGEVAKLPADAGVICTASLISPYRRDRACRALLPHSN 264
Db 67 VRHGLCRDLGFDADQRQENIRVGEVASLMDAGLIVLTAFTSPHRAERQLKRVGHDR 126

QY 265 FIEVFIDLPLKICEARDPKGLYKLTGKIGFTGIDDPYEP-----INGEIVI 314
Db 127 FIEIVNTPLAICEQRDPKGLYKKARAGELRNFTGIDAIYEAPDSPQVHLNGEQLV 182

RESULT 15
JW0087
adenylyl-sulfate kinase (EC 2.7.1.25) - human
N:Alternate names: adenosine 5'-phosphosulfate kinase; PAPS
N:Contains: adenylylsulfate kinase (EC 2.7.1.25); sulfate adenylyltransferase (EC 2.7.7.
C:Species: Homo sapiens (man)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: JW0087
R:Yanagisawa, K.; Sakakibara, Y.; Suiko, M.; Takami, Y.; Nakayama, T.; Nakajima, H.; Tak
Biosci. Biotechnol. Biochem. 62, 1037-1040, 1998
A:Title: cDNA cloning, expression, and characterization of the human bifunctional ATP su
A:Reference number: JW0087; MUID:98312048; PMID:9648242
A:Accession: JW0087
A:Molecule type: mRNA
A:Residues: 1-624 <VAN>
A:Cross-references: UNIPROT:O43252; GB:AF033026; NID:g3378100; PIDN:AAC28429.1; PID:g337
A:Experimental source: Brain
C:Function: <ASKF>
A:Description: as adenylylsulfate kinase catalyzes the phosphorylation of adenylylsulfat
C:Function: <SATF>
A:Description: as sulfate adenylyltransferase catalyzes the reaction of sulfate and ATP
C:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenylylsulfate
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loop;
F:52-215/Domain: adenylylsulfate kinase homology <ASK>
F:59-66/Region: nucleotide-binding motif A (P-loop)
```

```
F:226-620/Domain: sulfate adenylyltransferase homology <SAT>
F:133/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 27.1%; Score 495; DB 1; Length 624;
Best Local Similarity 46.1%; Pred. No. 1.4e-29;
Matches 106; Conservative 38; Mismatches 60; Indels 26; Gaps 6;

QY 125 EAPHSPVKEKPVMSN-----IGKSTNIIWHNCLIGSDRQKLLGQ-----KGVVVIITGLS 175
Db 2 EIPGSLCKVKVLSNNAQNWGMORATNTVYQAHVSRNKRQGVVTRGGFRGCTVWLITGLS 61

QY 176 GSGKSTLACALSRELHCRGHLYTVLVDGNLRHGLNRDLSPKAEDRAENIRVGEVAKLFA 235
Db 62 GAGKTTVSMALVEYLCHGIPCYTLGDGNIROGLNKNLGFSPEDREENVRRIRAEVAKLFA 121

QY 236 DAGVICTASLISPYRRDRDADR-----ALLPHSNFIEVFIDLPLKICEARDPKGLYKLR 290
Db 122 DAGLVCTISFTISPYTQDRNNARQIHEGASLP---FFEVFVDAPLHVCEQRDVKGLYKKAR 178

QY 291 TGKIKGFTGIDDPYEPPIINGEIVIKMKD---EECPSPKAMAKQVLCYLEE 337
Db 179 AGEIKGFTGIDSEYEKPEAPELVLTDSVDYND-----VQQVVELLOE 222
```

Search completed: September 9, 2005, 13:04:01
Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 9, 2005, 13:04:07 ; Search time 3067 Seconds
(without alignments)
4256.943 Million cell updates/sec

Title: US-10-829-432-4

Perfect score: 1826

Sequence: 1 RPFHINOTPLVTHQPPP.....PRMAKQVLCYLENGYLQA 343

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool_p/US10829432/runat_08092005_132223_25088/app_query.fasta_1.519
-DB=EST -Qfmt=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10829432 @CNG_1_1_5180 @runat_08092005_132223_25088 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1246	68.2	720	7	COS28976
2	1080.5	59.2	645	7	COS32501
3	944	51.7	915	9	CL981208
4	936	51.3	675	2	BE357876
5	926	50.7	691	2	BE355111
6	881.5	48.3	1075	7	CK205911
7	874	47.9	614	6	CA204452
8	865	47.4	617	6	CD863007
9	863.5	47.3	860	7	CR285854

10	837.5	45.9	673	6	CA182186
11	816.5	44.7	867	7	CO095086
12	816	44.7	1066	7	CNS0A87V
13	810.5	44.4	991	3	AY110087
14	801.5	43.9	659	6	CA900730
15	798	43.7	970	3	AY109629
16	797	43.6	555	4	B1427055
17	797	43.6	809	6	CA766480
18	793	43.4	696	6	CA085472
19	792.5	43.4	837	7	CO200697
20	789	43.2	807	7	CNI134308
21	775	42.4	797	7	CNI29765
22	774.5	42.4	1206	3	CNS0A2AD
23	773	42.3	754	6	CA921172
24	772	42.3	1090	5	BX838386
25	760.5	41.6	648	6	CD817873
26	757.5	41.5	1163	3	CNS091ZG
27	751	41.1	536	2	AW560397
28	748.5	41.0	1146	3	CNS0A2PA
29	748	41.0	525	6	CA502007
30	747	40.9	595	5	BU808718
31	744	40.7	872	7	CV281393
32	743	40.7	673	6	CD924390
33	738	40.4	496	6	CA497171
34	735	40.3	902	7	CV245034
35	733.5	40.2	748	4	BG126657
36	728	39.9	564	2	AW056154
37	728	39.9	957	7	CV260804
38	725	39.7	860	7	CF515151
39	723	39.6	700	1	AJ502331
40	717	39.3	803	7	CO476247
41	717	39.3	814	6	CA759382
42	714	39.1	1158	3	CNS092LM
43	712	39.0	772	2	BF631366
44	711	38.9	846	7	CF515003
45	710	38.9	617	5	BQ518805

ALIGNMENTS

RESULT 1
COS28976 3530.1_191_1_D07.Y.1_3530 - Full length cDNA library created by
LOCUS Invitrogen from multiple tissues Zea mays cDNA, mRNA sequence.
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

COS28976 3530.1_191_1_D07.Y.1_3530 - Full length cDNA library created by
Invitrogen from multiple tissues Zea mays cDNA, mRNA sequence.
COS28976.1 GI:50333850
EST.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 3530.1_191.1 row: D column: 07.
Location/Qualifiers
1..720
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="B73"
/db_xref="taxon:4577"
/tissue_type="multiple"

FEATURES

source
1..720
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="B73"
/db_xref="taxon:4577"
/tissue_type="multiple"

/dev_stage="varies by tissue"
 /lab_host="DH10B"
 /clone_lib="3530 - Full length cDNA library created by
 Invitrogen from multiple tissues"
 /note="Organ: silks, husks, ears, pollen, shoot tips,
 leaf, root tips, whole seed, embryo; Vector: pCMV-SPORT
 6.1; Site 1: EcoRV; Site 2: NotI; Maize Gene Discovery
 Project contracted with Invitrogen to produce a
 normalized, full length library in a pSPORT vector. This
 is a Gateway compatible vector, permitting clone movement
 to new vector backbones for expression in diverse host
 cells using recombination rather than restriction enzymes.
 Details of the vector and sequencing primers are available
 at ZmDB in the EST library description tables. poly(A)+
 mRNA was prepared by Invitrogen, and equimolar amounts of
 RNA from each of the 12 tissue samples were mixed together
 for selection of mRNA with a 5' cap. After synthesis of
 cDNA, a normalization step was conducted against the
 mixture of RNA sources. This step effected a 20X to 80X
 reduction in common transcript types. Tissues prepared: 1.
 just emerging silks; 2. inner husks from ears of sample
 #1; 3. 20 dap aleurone; 4. immature tassels, stages from
 1-2 mm to 1-2 cm; 5. 2 mm to 2 cm ears; 6. pollen; 7. 1 cm
 vegetative shoot tips from 15 day old seedlings; all
 leaves with an expanded or partially expanded sheath
 were removed; 8. mature leaf tissue; 9. 0.5 cm long root
 tips from 15 day old seedlings; 10. 10 dap whole seed; 11.
 12 dap endosperm and embryo; 12. 17 dap endosperm and
 embryo. All of the sequenced clones in project 3530 will
 be archived at the University of Arizona along with the
 Unigene clones from the Maize Gene Discovery EST
 sequencing projects. Clones can be ordered through the
 ZmDB web site or directly from the University of Arizona
 (<http://www.genome.arizona.edu/orders/>). High density
 filters containing over 18,000 clones can also be ordered
 from the University of Arizona."

ORIGIN

Alignment Scores:
 Pred. No.: 1,15e-87 Length: 720
 Score: 1246.00 Matches: 238
 Percent Similarity: 99.58% Conservative: 1
 Best Local Similarity: 99.17% Mismatches: 0
 Query Match: 68.24% Indels: 1
 DB: 7 Gaps: 0

US-10-829-432-4 (1-343) x COS28976 (1-720)

QY	28	SerGlnGlyGlnArgGlnGlyAsnThrLeuLeuSerProThrProThrLeuAlaValle	47
Db	2	AGCCAGGCGCAACGGCAAGCAACACCTCTCTCAGCCGACGCCGCTGCGGTCATC	61
QY	48	LeuValAsnProGlnArgAlaProProValLeuProGlyLeuThrProSerAspAlaPro	67
Db	62	CTCGTAATCCACAGCGCGCGCTCCCGTCCTCCAGCGCTCACCCTTAGCGATCGCCA	121
QY	68	LeuProAlaLeuValIleHisGlyLeuThrProArgSerSerHisSerAlaGlyLeu	87
Db	122	CTCCCGGCGCTGATCATCGCTCACTCCCGCTTCTCACACTCTCCGCGGTC	181
QY	88	AlaSerAspSerGlyArgArgGluGlyGluGlyArgGlyAlaArgThrHisCysHisArg	107
Db	182	GCCAGCATAGTGGGCGCGCGAGGGAGGGCGCGGTGCGGTACGCACTGCCACCGC	241
QY	108	GlyIleGlyArgTrpValArgArgArgArgAsnGlyAlaAlaPro-GlyGluAlaPr	127
Db	242	GGCATTTGGCGGTGGGTGGCGCGCGCGCGGAATGGAGACGCCACGGGGAGCCCC	301
QY	127	HisSerProVallySGLuLysProValMetSerAsnIleGlyLysSerThrAsnIleLe	147
Db	302	GCACAGCCAGTGAAGGAGAACCTGTATGTCAACATTTGGAAATCGACTAATATTTT	361
QY	147	utrHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuGlyGlnLysGlyCy	167

Db	362	ATGCCACAATTCCTTGATGGACAATCTGTAGACAGAAATTCGTGGCAGAAAAGCGTCG	421
QY	167	sValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSe	187
Db	422	TGTCGTATGGATAACAGGACTCAGTGGTTAGGGAAAAGTACTCTTTCATGTGCACTGAG	481
QY	187	rArgGluLeuHisCysArgGlyHisLeuThrTyValLeuAspGlyAspAsnLeuArgHi	207
Db	482	TCGTGAGTTGCATTGCAGAGGCCACCTCACGTATGTACTTGTATGGTGACCACTCAGACA	541
QY	207	sGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgArgVa	227
Db	542	TGGCCTTAATAGAGATTAAAGCTTTAAGGCGCAGAGACCGTCAGAAAATATACGAAGAGT	601
QY	227	lGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSe	247
Db	602	TGGTGAAGTGGCAAGCTTTTCGCTGATGCTGGTGTCTATGTCATTGCTAGCTTGATATC	661
QY	247	rProTyArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSerAsnPheIle	266
Db	662	TCCATACAGGAGAGATCGTGATGATGCGCGTCTTACTTCCACATTTAACTTTATT	719

RESULT 2
 COS32501
 LOCUS
 DEFINITION 3530.1.213.1.H05.Y.1.3530 - Full length cDNA library created by
 Invitrogen from multiple tissues Zea mays cDNA, mRNA sequence.
 ACCESSION COS32501
 VERSION COS32501.1 GI:50337375
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (Bases 1 to 645)
 AUTHORS Walbot,V.
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 UNIVERSITY
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 3530.1.213.1 row: H column: 05.
 Location/Qualifiers
 1..645
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /tissue_type="multiple"
 /dev_stage="varies by tissue"
 /lab_host="DH10B"
 /clone_lib="3530 - Full length cDNA library created by
 Invitrogen from multiple tissues"
 /note="Organ: silks, husks, ears, pollen, shoot tips,
 leaf, root tips, whole seed, embryo; Vector: pCMV-SPORT
 6.1; Site 1: EcoRV; Site 2: NotI; Maize Gene Discovery
 Project contracted with Invitrogen to produce a
 normalized, full length library in a pSPORT vector. This
 is a Gateway compatible vector, permitting clone movement
 to new vector backbones for expression in diverse host
 cells using recombination rather than restriction enzymes.
 Details of the vector and sequencing primers are available
 at ZmDB in the EST library description tables. poly(A)+
 mRNA was prepared by Invitrogen, and equimolar amounts of
 RNA from each of the 12 tissue samples were mixed together
 for selection of mRNA with a 5' cap. After synthesis of

QY 109 leGlyArgTrrpValArgArgArgArgAsnGlyAlaala---ProGlyGluAlaProH 128
 Db 184 --AGCAGCTGGACAGGAAGCTGCCGCTCGAGCGCGCCCTGCGCGCGCGCGCG 241
 QY 128 isSer-----ProValLysGluLysProValMetSerAsnI 140
 Db 242 TCGACGGGAAGACGACGACGATGTTGCCAGTTTAAACAAGCTCATGACCTCAACTG 301
 QY 140 leGlyLysSerThrAsnleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnL 160
 Db 302 TCGGTAATACAAACATCCGATGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
 QY 160 ysLeuLeuGlyGlnLysGlyCysValValTrrpIleThrGlyLeuSerGlySerGlyLys 180
 Db 362 AGTTGCTGAACCAAGAGGCTGTGTGTTGGATCACTGCTGCTGCTGCTGCTGCTGCTG 421
 QY 180 exThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrrValL 200
 Db 422 GCACCTTGTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
 QY 200 euAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelysAlaGluAspA 220
 Db 482 TTGACGGCGACAACTCTCCGGCAGCGCTGAAACAAAGATCTCAGCTTCAAGCCCAAGGATC 541
 QY 220 tGAlaGluAsnleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValI 240
 Db 542 GTGCCGAAATATACGAGCTTGAGAAAGTGGCAAGCTGTTGCAATGCTGATGTA 601
 QY 240 leCysleAlaSerLeuLeuSerProTrrArgArgAspAlaCysArgAlaLeuL 260
 Db 602 TCTGATTACTAGTTTGTATACCTATACCTATACCTATACCTATACCTATACCTATAC 661
 QY 260 euProHisSerAsnPhelleGluValPheIleAspLeuProLeuLysleCysGluAlaA 280
 Db 662 TACCAATTTCTTGGTTCATAGGCTGTTCTGTAATGTCCCACTTGAAGTGTGAAGAA 721
 QY 280 tGAspProLysGlyLeuTrrLysLeuAlaAtqThrGlyLysleLysGlyPheThrGlyV 300
 Db 722 GGGATCCAAAGGCTGTACAGCTTGTCTGTCGGCGGCAAAATCAAGGCTTACGGAA 781
 QY 300 leAspAspProTrrGluProPheIleAsnGlyGluIleValleLysMetLysAspGluG 320
 Db 782 TAGATGATCTTATGAACACCTTCAGATTGTGAGATTGTGATACAGTGCACAAAGTTGGG 841
 QY 320 luCysProSerProLysAlaMetAlaLysGlnValLeuCysTrrLeuGluGluAsnGlyT 340
 Db 842 ACTGCCCTTCACTAAATCAATGGCTGATCAAGTAGTGTATCTTGAAGCCCAATGGAT 901
 QY 340 yrLeuGln 342
 Db 902 TCTTTCAG 909

RESULT 4
 BE357876
 LOCUS DG1.22.A06.g1.A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
 DEFINITION

BE357876
 BE357876.1 GI:9299433
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Sorghum bicolor (sorghum)
 Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE
 AUTHORS
 Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and
 Pratt,L.H.

TITLE
 JOURNAL
 COMMENT
 An EST database from Sorghum: dark-grown seedlings
 Unpublished (2000)
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1850
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions
 below Phred quality 16. The threshold for highest quality sequence
 is 20.

Seq primer: PolyTMix
 High quality sequence start: 30
 High quality sequence stop: 638
 POLYA-No.

FEATURES

source

1. .675
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /db_xref="taxon:4558"
 /clone_lib="Dark Grown 1 (DG1)"
 /note="Organ: 5-day-old dark-grown seedlings; Vector:
 Lambda Zap; Site 1: XhoI; Site 2: EcoRI; The library was
 made from poly-A⁺ RNA in the cloning vector lambda ZAP II.
 Clones to be sequenced were prepared by mass excision."

ORIGIN

Alignment Scores:
 Pred. No.: 2.16e-63 Length: 675
 Score: 936.00 Matches: 178
 Percent Similarity: 96.81% Conservative: 4
 Best Local Similarity: 94.68% Mismatches: 6
 Query Match: 51.26% Indels: 0
 DB: 2 Gaps: 0
 US-10-829-432-4 (1-343) x BE357876 (1-675)

QY 156 SerAspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrrpIleThrGlyLeuSer 175
 Db 2 TCTGATACAGAAAGTTGCTCGGCAAAAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 61
 QY 176 GlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHis 195
 Db 62 GTTCAGGGAAGAGTACTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
 QY 196 LeuThrTrrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhe 215
 Db 122 ATCAGTATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTT 181
 QY 216 LysAlaGluAspArgAlaGluAsnIleArgValGlyGluValAlaLysLeuPheAla 235
 Db 182 AAGCAGAAAGACCGTGCAGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGGC 241
 QY 236 AspAlaGlyValIleCysIleAlaSerLeuIleSerProTrrArgArgAspAspAla 255
 Db 242 GATGCTGGTATCATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
 QY 256 CysArgAlaLeuLeuProHisSerAsnPhelleGluValPheIleAspLeuProLeuLys 275
 Db 302 TCCGCTGCTACTTCCAGATTCTAACTTTATTGAAGTATTATTGATGATTTGCCCTAAAA 361
 QY 276 IleCysGluAlaArgAspProLysGlyLeuTrrLysLeuAlaArgThrGlyLysleLys 295
 Db 362 ATTTGTGAAGCTCGTGATCTTAAAGGCTGTGACAGGCTTGCAGGACAGGAAGATTAAA 421
 QY 296 GlyPheThrGlyIleAspAspProTrrGluProPheIleAsnGlyGluIleValleLys 315
 Db 422 GGTTCATCGAATGATGATCCATACCAACCGCCAGTTAATGTTGATGATGATGATGATG 481
 QY 316 MetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTrrLeu 335
 Db 482 ATGAAGATGGGAATGCCCTTCCACCAAGCAATGCGGAAGCAAGTTCTATCATACCTT 541
 QY 336 GluGluAsnGlyTrrLeuGlnAla 343
 Db 542 GAAGAGAACGGATATTTCGAAGCT 565

having emerged from the leaf (visible) . First strand synthesis in this library was done in the presence of methylated dCTP thereby protecting from internal cleavage with NotI."

ORIGIN

Alignment Scores:			
Pred. No.:	7,69e-59	Length:	1075
Score:	881.50	Matches:	188
Percent Similarity:	78.2%	Conservative:	21
Best Local Similarity:	70.41%	Mismatches:	41
Query Match:	48.27%	Indels:	21
DB:	7	Gaps:	2
US-10-829-432-4 (1-343) x CK205911 (1-1075)			
QY	77	ThrProArgSerSerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgGluGly	96
Db	1073	AGCCCCCGTGGCGCTGGCGCTGGCGCTCCGTTGGCCCACTTGGCCGGAACGGGGGTGC	1014
QY	97	GluGlyArgGlyAlaArgThrHisCysHisArgGlyIleGlyArgTyrValArgArg	116
Db	1013	CGCGAACCAGTGAACACCCCTCTC-----	987
QY	117	ArgArgAsnGlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProVal	136
Db	986	---AGGACGGGGCCCTTCG-----AAGCCCCAAGTGGAAAAAGCCTTTA	945
QY	137	MetSerAsnIleGlyLysSerThrAsnIleLeuTyrHisAsnCysLeuIleGlyGlnSer	156
Db	944	AT-TCCAACTATGGGAA-TCCAAATTAATTTTATGGCATGACTGCCCAATTTGGACAAAT	887
QY	157	AspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTyrIleThrGlyLysSerGly	176
Db	886	GAGCAGACAGAAATGCTGGGGCCAAAGAGATTTTCATTGGATCACAGGACTCAGCGGT	827
QY	177	SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu	196
Db	826	TCAGGAAAAAGTACCTT-GCCTGTGCACCTGAT-CGGGGATTACACTACAGAGCCACCAC	769
QY	197	ThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelys	216
Db	768	ACGTATGCTTGAATGGTGACAACTCAGACATGCGCTCAATCGAGATCTAAGCTTCAAG	709
QY	217	AlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAsp	236
Db	708	GCAGAGACCGTAGAGAAATATACAGAGTTGGTGAAGTGGCAAGCTTTTTCAGAT	649
QY	237	AlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAlaCys	256
Db	648	GCTGTACCATATGCATTGCTAGTTTGATATCTCCATACAGGAGAGCCGTGATGCATGC	589
QY	257	ArgAlaLeuLeuProHisSerAsnPhelleGluValPheIleAspLeuProLeuLysIle	276
Db	588	CGAGCTCTACTTCCAGATTTCTAGATTTATTTGAAGTATTTATGGATTTGCCACTAGAATA	529
QY	277	CysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGly	296
Db	528	TGTGAAGCTCGTGATCCTTAAGGGTTATACAAGTTTGCACGACAGGAAGATTAAAGG	469
QY	297	PheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMet	316
Db	468	TTCCCGGAGTTGATGATCATCACAAATCACCCAGTGAATAGTAGATGATTAATTAAGATG	409
QY	317	LysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuLysCysTyrLeuGlu	336
Db	408	GAAGGTGGGAATGCCCTTTCACCGAAAGCAATGGGCCCAAGTTCTATCCTTACCTTGAG	349
QY	337	GluAsnGlyTyrLeuGlnAla	343
Db	348	AAGATGGATATTTCAGGCT	328

CA204452	CA204452	614 bp	mRNA	linear	EST 25-SEP-2003
LOCUS	SCAGFL1086H06.g FL1 Saccharum officinarum cDNA clone SCAGFL1086H06				
DEFINITION	5', mRNA sequence.				
ACCESSION	CA204452				
VERSION	CA204452.1	GI:35241732			
KEYWORDS	EST.				
SOURCE	Saccharum officinarum				
ORGANISM	Saccharum officinarum				
REFERENCE	1 (bases 1 to 614)				
AUTHORS	Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.				
TITLE	The libraries that made SUCEST				
JOURNAL	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)				
COMMENT	Contact: Arruda P Centro de Biologia Molecular e Engenharia Genetica Universidade Estadual de Campinas Caixa Postal 6010, 13083-970, Campinas SP, Brazil Tel: 55 19 3788 1137 Fax: 55 19 3788 1089 Email: parruda@unicamp.br Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bccccenter.fcav.unesp.br Plate: 086 ROW: H column: 06 Seq primer: T7 Promoter Primer. Location/Qualifiers 1. .614 /organism="Saccharum officinarum" /mol_type="mRNA" /db_xref="taxon:4547" /clone="SCAGFL1086H06" /lab_host="DH10B" /clone_lib="FL1" /note="Organ: Inflorescence at beginning of development (1cm-long); Vector: pSport1, Site_1: SalI; Site_2: NotI; An unidirectional cDNA library generated from [Inflorescence at beginning of development (1cm-long)] cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.lad.ic.unicamp.br/public"				
FEATURES	source				
ORIGIN					
Alignment Scores:					
Pred. No.:	1.38e-58	Length:	614		
Score:	874.00	Matches:	178		
Percent Similarity:	89.16%	Conservative:	3		
Best Local Similarity:	87.68%	Mismatches:	14		
Query Match:	47.86%	Indels:	8		
DB:	6	Gaps:	3		
US-10-829-432-4 (1-343) x CA204452 (1-614)					
QY	45	AlaValIleLeuValAsnProGlnArgAlaProProValLeu-ProGlyLeuThrProSe	64		
Db	18	GCCGTCGCGCCCGTAATCCATAGCGCGCGGTCCGTCCTCCGCC-----ACCCCCA	71		
QY	64	rAspAlaProLeuProAlaLeuValIleHisGlyLeuThrProArgSerHisSerSe	84		
Db	72	CGATGCTCCGTGCGGGCTCGTGATCCATGGCTTCGCTCCGCTTCTCACAGCTCC	131		
QY	84	rAlaGlyLeuAlaSerAspSerGlyArgArgGluGlyGlu-----GlyArgGlyAl	101		
Db	132	CGCGGGCTCCCGAGCGGTGTGGCGCGCGAGGGGGAGGAGAGCGCGCGGTGC	191		
QY	101	aArgThrHisCysHisArgGlyIleGlyArgTyrValArgArgArgArgAsnGlyAl	121		

```

Db      192 GCGTGGCACTGCCCGGCAATGGCGGTGGATG---CGGTGGCGGGAATGGAGC 248
Qy      121 aalaPro-GlyGluAlaProHisSerProValIysGluLysProValMetSerAsnIleG 141
Db      249 AGCGCCACGGGGAGGACACACCGCCAGTGAAGAGAGGCTGTAAATGTCAAAATGTG 308
Qy      141 lYlYsSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysL 161
Db      309 GGAATCAACTAATATTTATGGCATAATTTGCCGATTTGGCAATCTGATAGACAGAAGT 368
Qy      161 euLeuGlyClnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerT 181
Db      369 TGCTGGGACAAAGAGCTGTGCTGTGGATAACAGGACTCAGTGGTTCAGGGAAGATA 428
Qy      181 hcLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpValLeuA 201
Db      429 CTCCTGTCATGTGCTGAGTGTGAGTGTGATATAGAGCCACCTCCATATGTACTTG 488
Qy      201 spGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelYsAlaGluAspArgA 221
Db      489 ATGGTGACAACTTAGACATGCGCTAAATCGAGATTAAAGCTTTAAGGCAGAGACCGTG 548
Qy      221 lsGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleC 241
Db      549 CAGAAATATACGAAGAGTTGTTGAAGTGGCAAGCTTTTCCCGATGCTGTATCATAT 608
Qy      241 yaile 242
Db      609 GCATT 613

```

```

RESULT 8
CD863007 617 bp mRNA linear EST 11-JUL-2003
LOCUS AZ01.105F21F010130 AZ01 Triticum aestivum cDNA clone AZ01105F21,
DEFINITION mRNA sequence.
ACCESSION CD863007
VERSION CD863007.1 GI:32546823
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 617)
Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).
Location/Qualifiers
1..617
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZ01105F21"
/tissue_type="leaf"
/clone_lib="AZ01"

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).
Location/Qualifiers
1..617
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZ01105F21"
/tissue_type="leaf"
/clone_lib="AZ01"

```

FEATURES

```

source
1..617
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZ01105F21"
/tissue_type="leaf"
/clone_lib="AZ01"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 7,07e-58 Length: 617
Score: 865.00 Matches: 163
Percent Similarity: 93.51% Conservative: 10
Best Local Similarity: 88.11% Mismatches: 12
Query Match: 47.37% Indels: 0

```

```

DB: 6 Gaps: 0
US-10-829-432-4 (1-343) x CD863007 (1-617)
Qy 159 GlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGly 178
Db 1 CAGAAATTCCTGGGGCAAAAAGAGTGTGTATATGGATCACAGGACTCAGCGGTTCAGGG 60
Qy 179 LysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrp 198
Db 61 AAAGTAGCCGTGCTGTGCACTGAGTCGGAAATTACACTACAGAGGCCACACACGATAT 120
Qy 199 ValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelYsAlaGlu 218
Db 121 GTTCTTGATGTGTGCAACCTCAGACATGGCTTAATCGAGATCTTAAGCTTCAAGGCAGAA 180
Qy 219 AspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGly 238
Db 181 GACCGTACAGAAATATATACGAAGAGTTGGAGAAGTGGCAAGCTTTTTCAGATGCTGGT 240
Qy 239 ValIleCysIleAlaSerLeuIleSerProTrpArgArgAspArgAspAlaCysArgAla 258
Db 241 ACCATATGCATGCTAGTTTGTATATCTCCATACAGAGAGACCGGTGATGATGCCAGCT 300
Qy 259 LeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGlu 278
Db 301 CTACTTCCAGATCTTAGATTTTATGAAGTATTTATGGATTTGCCACTAGATAATATGTCAA 360
Qy 279 AlaArgAspProLysGlyLeuTrpLysLeuAlaArgThrGlyLysIleGlyGlyPheThr 298
Db 361 GCTCGTGATCTCTAAAGGGTTATACAAGCTTTCACGCACAGGAAAGATTAAAGGGTTTACC 420
Qy 299 GlyIleAspAspProTrpGluProPheIleAsnGlyGluIleValIleLysMetLysAsp 318
Db 421 GGAGTGTATGATTCATACAGATCACCAGTGAATAGTAGATAGTAGTAATTAAGATGAAGGT 480
Qy 319 GluGluCysProSerProLysAlaMetAlaLysGlnValLeuLysCysTrpLeuGluGluAsn 338
Db 481 GGGGAATGCCCTTCAACGAGGCAATGCCCGAGCAAGTTCTGTCTTACCTTGAGAAGAAGC 540
Qy 339 GlyTrpLeuGlnAla 343
Db 541 GGATATTTTGCAGGCT 555

```

RESULT 9

```

CD828584
LOCUS CR285854
DEFINITION CR285854 Oryza sativa library (Han B) Oryza sativa cDNA clone
y625e05p5, mRNA sequence.
ACCESSION CR285854
VERSION CR285854.1 GI:44672420
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa

```

ORGANISM

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 860)

```

REFERENCE

```

AUTHORS

```

```

Hao, B., Feng, Q., Huang, Y.C., Ying, K., Li, Y., Guan, J.P., Zhu, J.J.,
Zhao, Q., Hu, X., Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L.,
Weng, Q., Zhang, L., Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.,
Zhang, Y.J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X.,
Zhang, L., Lan, L.F., Chen, W., Wu, S.A. and Xue, Y.B.
Rice cDNA EST clone
Unpublished (2003)
Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences
500# Cao Bao Road, Shanghai 200233, China
Email: bhan@ncgr.ac.cn
Clone requests: bhan@ncgr.ac.cn
This is rice cDNA est clone
Web site: http://www.ncgr.ac.cn.

```

TITLE

```

JOURNAL
COMMENT

```

CAD

RESULT 10					
CA182186	CA182186	673 bp	mRNA	linear	EST 24-SEP-2000
LOCUS	SCBGST3108C02.g	STP3	Saccharum officinarum	cDNA clone	SCBGST3108C02
DEFINITION	5'.. mRNA sequence.				
ACCESSION	CA182186				
VERSION	CA182186.1	GI:35117704			
KEYWORDS	EST.				
SOURCE	Saccharum officinarum				
ORGANISM	Saccharum officinarum				

```

cores:      1.14e-55      Length:      673
            837.50       Matches:      186
            75.30%      Conservative:    3
            74.10%      Mismatch:         27
            45.87%      Indels:          36
            6           Gaps:            5

32-4 (1-343) x CA182186 (1-673)

16 ThrGlnProProSerProAlaProGlyProAlaSerGlnGlyGlnArgGlnGlyAsn 35
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
12 AGCCAGCAACACCTCTTCGGTCCGACGCTACGCCGGAGCGGCACCGC----- 62
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

136 ThrLeuLeuSerProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaPro 55
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
63 -----GCCGTCGGCCCCGTAATCCATAGCGCGCGCT 95
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

56 ProValLeuProGlyLeuThr-ProSerAspAlaProLeuProAlaLeuValIleHisG1 75
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
96 CCGCGTCCTCCCAACCCCAACCCCAACGATGCGCGCTGCTGGAGCTCGTGATCCAT-- 153
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

75 yLeuThrProArgSerSerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgG1 95
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
54 -----GGCTCCCAAGCGGTGGTGGCGCGCCGA 182
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

95 uGly-----GluGlyArgGlyAlaArgThrHisCysHisArgGlyIleGlyArgTr 112
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
83 GGGGAGGAGGAGGAGGCGCGGTGCGCGTGCACCTGCCACCGCGCATTTGGCGGTG 242
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

12 pValArgArgArgArgAsnGlyAlaAlaPro-GlyGluAlaProHisSerProVal1 132
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```


Db 243 G-----CGCGGAGTGGAGCAGCGCCACCGGGAG-----GGCCCAAGTGA 281

Qy 132 ysGluysProValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysL 152

Db 282 AGGAGAGCGCTGTAATGTCAAACATGGGAATCAACTAATATTTATGGCATAATGCC 341

Qy 152 euIleGlyClnSerAspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleT 172

Db 342 CGAATGGACAAATCTCATCACAGAAGTTGCTGGGACAAAAGGCTT-GTCGTGTGGATA 400

Qy 172 hrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisC 192

Db 401 CAGGACTCAGTGGTTCAGGGAAGTACTCTTGCATGTGCACTGAGTCATGAGTTGCATT 460

Qy 192 ysArgGlyHisLeuThrTrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgA 212

Db 461 ATAGAGGCCACCTCACATATGCTTGTGATGTGACAACTTAGACATGGCCTAATCGAG 520

Qy 212 spLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaL 232

Db 521 ATTTAAGTTTTAAAGCAGAGACCGTCGAGAAAATATACGACGAGTTGGTGAAGTGCCAA 580

Qy 232 ysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTrpArgArgA 252

Db 581 AGCTTTTTCGCCATGCTGGTATCATATGCAATTCATGCTTAGCTTGATATCTCCATACAGGAG 640

Qy 252 spArgAspAlaCysArgAlaLeuLeuPro 261

Db 641 ATCGTATGATGCCCGGCTCTACTTCCA 669

RESULT 11

LOCUS CO095086 867 bp mRNA linear EST 16-JUN-2004

DEFINITION GR_Eal7119.r GR_Ea Gossypium raimondii cDNA clone GR_Eal7119 3', mRNA sequence.

ACCESSION CO095086

VERSION CO095086.1 GI:48793772

KEYWORDS EST.

SOURCE Gossypium raimondii

ORGANISM Gossypium raimondii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

1 (bases 1 to 867)

Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,

Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and

Wing, R.A.

Global assembly of Cotton ESTs

Unpublished (2004)

Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: <http://genome.arizona.edu>

Plate: 17 row: I column: 19.

Location/Qualifiers

1..867

/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_Eal7119"

/tissue_type="whole seedlings"

/dev_stage="first true leaves"

/lab_host="DH10B"

/clone_lib="GR_Ea"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:

EcoRV; Library made by Invitrogen with RNA supplied by

Wendle lab. Directional clones into NotI-EV. Colonies

plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:

Pred. No.: 7.17e-54 Length: 867

Score: 816.50 Matches: 152

Percent Similarity: 85.44% Conservative: 24

Best Local Similarity: 73.79% Mismatches: 29

Query Match: 44.72% Indels: 1

DB: 7 Gaps: 1

US-10-829-432-4 (1-343) x CO095086 (1-867)

Qy 137 MetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyClnSer 156

Db 45 ATGTCGACTATGGGAAATTCGACAAATATATTTTGGCAAGAAATCCGCTGTGGGAAGCTT 104

Qy 157 AspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGly 176

Db 105 GAGAGCGGAGCTTACTTACCAAAAGGGTTGTGTATGGATCCTGGTCTCAGCGGA 164

Qy 177 SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu 196

Db 165 TCAGGTAAAGCACACTTGCAATGTTCTTAAGTTGGGAACCTTCATACAGAGGCAAGCTA 224

Qy 197 ThrTrpValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLys 216

Db 225 TCTTACATACTTGTATGGGACACCGTTGCGCAGGATTTAAACAAGGATCTTGGTTTCAAG 284

Qy 217 AlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAsp 236

Db 285 GCTGAGGATCGACAGAAAATATTCGCAGGGTTGGTGAAGTGGCAAGCTCTTTGCAGAT 344

Qy 237 AlaGlyValIleCysIleAlaSerLeuIleSerProTrpArgArgAspArgAspAlaCys 256

Db 345 GCTGTTTAAATCTGCATTCGCCAGTCTAATATCTCCATATGGGAAAGACCGTGATCATGC 404

Qy 257 ArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIle 276

Db 405 CGAGCAATGTTGCCGGATGCT---TTTATTGAGGTTTTCATGAACATGCCCTAGCATTA 461

Qy 277 CysGluAlaArgAspProLysGlyLeuTrpLysLeuAlaArgThrGlyLysIleLysGly 296

Db 462 TGTGAAGCGCAGATCCGAAAGGCTTTTACAAAGCTTGCCCGTGTGGAAAGATTAAAGGT 521

Qy 297 PheThrGlyIleAspAspProTrpArgProLysGlyGluValIleLysMet 316

Db 522 TTTACTGTCATAGATGATCCTTACGAACACCTTTTGAACCTGTGAGATAGAATAAATCAG 581

Qy 317 LysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTrpLeuGlu 336

Db 582 AAAGATGGAGTTTGTCCACACCTAGTCCCATGCTGGGGAAGTAATTAATTACTTTGGAG 641

Qy 337 GluAsnGlyTrpLeuGln 342

Db 642 GACAAAGGATATCTGCAG 659

RESULT 12

LOCUS CNS0A87V

DEFINITION

CNS0A87V 1066 bp mRNA linear HTC 06-FEB-2004

Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSUS138ZF03 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION BX821394

VERSION BX821394.1 GI:42467158

KEYWORDS HTC; GSLT cDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1066)

Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,

Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,

Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:

TITLE

A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
 Unpublished
 2 (bases 1 to 1066)
 Genoscope.
 Direct Submission
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
 http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
 length
 http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
 FEATURES
 source
 1..1066
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSUFSILU382F03"
 /tissue_type="Silique"
 /plasmid="pCMVSPORT_6"
 1..1066
 /gene="At2g14750"

gene

ORIGIN

Alignment Scores:
 Pred. No.: 1.04e-53 Length: 1066
 Score: 816.00 Matches: 159
 Percent Similarity: 77.8% Conservative: 31
 Best Local Similarity: 65.16% Mismatches: 44
 Query Match: 44.69% Indels: 10
 DB: 3 Gaps: 3

US-10-829-432-4 (1-343) x CNS0A87V (1-1066)

Qy 100 GlyAlaArgThrHisCysHisArgGlyLeGlyArgTrpValArgArgArgAsn 119
 Db 204 GGATCTCAAACTCTGAGTCATAACAAAATGGATCTATTCTGAGGTTAAATCCATTAAAC 263
 Qy 120 GlyAlaAlaProGlyGluAlaProHisSerProValHisGlyLeuValMetSerAsn 139
 Db 264 GGT-----CACACGGGACAAAGCAAGGACCT---TTGTCTACG 299
 Qy 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuLeuGlyGlnSerAspArgGln 159
 Db 300 CTCGGAAACTCGACAAATATAAAGTGGCATGAATCTCTGTGAGAAAGTTGATAGACAG 359
 Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
 Db 360 AGATTGCTTGATCAGAAAGGAGTGTGATTTGGTCCACCGGCTTTAGTGGTTGAGGAAG 419
 Qy 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
 Db 420 AGTACTTTGGCTTTGGCTTTGAATCAGATGTTGATCAAAAGGGGAAGCTTTGTTATATT 479
 Qy 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
 Db 480 CTTGATGTTGATTAATGTTAGGATGGCTTAAACCCGTGATCTTAGCTTTAAAGCTGAGGAT 539
 Qy 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
 Db 540 CGTCAGAGATATTCTGTAGATTGGAGGTTGCTTAAGCTTTTTCGGATGCTGGATA 599

Qy 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspAlaCysArgAlaLeu 259
 Db 600 ATCTGCATTGGAGCTTTGATATCTCTATAGAACAGATAGGACGCTTGTGGAAGTTTG 659
 Qy 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
 Db 660 CTCGCCGAGGAGATTTTGTGAGGTTCATGATGTACCGCTTAGTGTGAGGCG 719
 Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
 Db 720 AGGATCCAAAGGCTCTTTACAAGCTCTCGTCGACGAAAGATCAAAAGGTTTACCCGG 779
 Qy 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
 Db 780 ATCGATGACCTTAGGAGCCACCATTTGAATGCCGAG-----ATTCTCTAGGACGTGA 833
 Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
 Db 834 GGAGGAATCTTCCTCATCGAAATGCGGAAAGGTCGTGCGATACTTAGATAACAAGGT 893
 Qy 340 TyrLeuGlnAla 343
 Db 894 TATCTTCAAGCA 905

RESULT 13
 AY110087
 LOCUS
 DEFINITION Zea mays CL1019_1 mRNA sequence.
 ACCESSION AY110087
 VERSION AY110087.1 GI:21214175
 KEYWORDS HTC.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 991)
 AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
 Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
 TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
 Overgo Probes
 JOURNAL Unpublished (2002)
 REFERENCE 2 (bases 1 to 991)
 AUTHORS Coe,E.H.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 COMMENT If you are interested in getting corresponding physical clones,
 these are publicly available from ZmDB and may be found by BLAST
 searching at MSI, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat
 Schnable, Iowa State, then clones may be requested from ZmDB:
 www.zmdb.iastate.edu.
 FEATURES
 source
 1..991
 Location/Qualifiers
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="maizdb:629888"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed DuPont contigs; this resource was
 assembled by DuPont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

ORIGIN
 Alignment Scores:
 Pred. No.: 2.54e-53 Length: 991
 Score: 810.50 Matches: 162

Db	420	CGATTTCATGTTACTGATAGAGAAATTTTATCGAGGTTTTCATGACATGCCTCTAGAAATG	479
QY	277	CysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGly	296
Db	480	TGTGAGCGACGAGATCAAAAGCGCTCTATAGCTTGCTCGTGGGGGAAAATCAAGGT	539
QY	297	PheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMet	316
Db	540	TTTACCGGCATTCATGACCCATATGACCAACCCATAAATTCGAGAGTTGAGATAAAGCAG	599
QY	317	LysAspGlu---GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeu	335
Db	600	GAGGATAGTGGCGTGTGTTCACACACCACTTTGATGGCTGCACAAAGTAGTTACTTACTTG	659
RESULT 15			
AY109629			
LOCUS	AY109629	Zea mays CL95_-2 mRNA sequence.	
DEFINITION	Zea mays CL95_-2 mRNA sequence.		
ACCESSION	AY109629		
VERSION	AY109629.1	GI:21213423	
KEYWORDS	HTC.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	1	(bases 1 to 970)	
AUTHORS	Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.		
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes		
JOURNAL	Unpublished (2002)		
REFERENCE	2	(bases 1 to 970)	
AUTHORS	Coe, E.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.		
FEATURES			
source	Location/Qualifiers		
	1..970		
	/organism="Zea mays"		
	/mol_type="mRNA"		
	/db_xref="MaizeDB:633120"		
	/db_xref="taxon:4577"		
	/clone_lib="Maize Mapping Project/DuPont Consensus Library"		
	/notes="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,36e-52	Length:	970
Score:	798.00	Matches:	170
Percent Similarity:	68.37%	Conservative:	31
Best Local Similarity:	57.82%	Mismatches:	58
Query Match:	43.70%	Indels:	36
DB:	3	Gaps:	6
US-10-829-432-4 (1-343) x AY109629 (1-970)			
QY	67	ProLeuProLalaValIleHisGlyLeuThrProArgSerHisSerAlaGly	96

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2005, 11:36:00 ; Search time 120 seconds
(without alignments)
1463.693 Million cell updates/sec

Title: US-10-829-432-4
Perfect score: 1826
Sequence: 1 RPFHFINTQTEPLVTHTOQP.....PRMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03: *
1: uniprot_spot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	873.5	47.8	345	Q6ZL22	Q6ZL22 oryza sativ
2	834	45.7	208	Q9SRW7	Q9SRW7 arabidopsis
3	833	45.6	312	KAP1_CATRO	Q49204 catharanthu
4	816	44.7	276	KAP1_ARATH	Q43295 arabidopsis
5	816	44.7	276	Q8LES2	Q8LES2 arabidopsis
6	812.5	44.5	288	Q9SE92	Q9SE92 zea mays (m
7	777.5	42.6	293	KAP2_ARATH	Q49196 arabidopsis
8	757.5	41.5	310	Q84JF0	Q84JF0 arabidopsis
9	755	41.3	290	Q9FXJ1	Q9FXJ1 arabidopsis
10	752.5	41.2	305	Q8LP64	Q8LP64 arabidopsis
11	632	34.6	152	Q9SQR9	Q9SQR9 arabidopsis
12	571	31.3	200	CYSC_CLOAB	Q97MT8 clostridium
13	563	30.8	618	Q9NDP8	Q9NDP8 ciona intes
14	555	30.4	199	CYC2_BACSU	Q06735 bacillus su
15	555	30.4	230	Q73WR1	Q73WR1 mycobacteri
16	550.5	30.1	201	Q8DGK5	Q8DGK5 synechococc
17	543.5	29.8	199	Q8EQN3	Q8EQN3 oceanobacill
18	538.5	29.5	207	CYSC_LACPL	Q98X60 lactobacill
19	534.5	29.3	202	CYC1_BACHD	Q9KCT0 bacillus ha
20	528.5	28.9	233	Q7XBB6	Q7XBB6 porphyra pu
21	527.5	28.9	614	Q842M0	Q842M0 rhodococcus
22	524.5	28.7	199	Q8CR04	Q8CR04 staphylococ
23	524	28.7	271	Q81351	Q81351 enteromorph
24	522.5	28.6	208	Q6LM71	Q6LM71 photobacter
25	521.5	28.6	197	Q63DV7	Q63DV7 bacillus ce
26	520.5	28.5	197	Q73B74	Q73B74 bacillus ce
27	519.5	28.5	197	Q6HLD2	Q6HLD2 bacillus th
28	519.5	28.5	208	Q7V3B6	Q7V3B6 prochloroc
29	519	28.4	205	Q6DIA7	Q6DIA7 erwinia car
30	514.5	28.2	197	Q81T47	Q81T47 bacillus an
31	514.5	28.2	214	Q7U3Z4	Q7U3Z4 synechococc

RESULT 1																				
Q6ZL22																				
ID	Q6ZL22	PRELIMINARY;	PRT;	345	AA.															
AC	Q6ZL22;																			
DT	05-JUL-2004	(Tremblrel. 27, Created)																		
DT	05-JUL-2004	(Tremblrel. 27, Last sequence update)																		
DT	05-JUL-2004	(Tremblrel. 27, Last annotation update)																		
DE	Putative adenosine-5'-phosphosulfate kinase.																			
GN	Name=OJ1699_E05.15;																			
OS	Oryza sativa (japonica cultivar-group).																			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;																			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;																			
OC	Ehrhartoideae; Oryzeae; Oryza.																			
OX	NCBI_TaxID=39947;																			
RN	[1]																			
RP	SEQUENCE FROM N.A.																			
RA	Sasaki T., Matsumoto T., Yamamoto K.;																			
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.																			
DR	EMBL;	AP003845;	BAC83149.1;																	
DR	GO;	GO:0005524;	F:ATP binding;	IEA.																
DR	GO;	GO:0016301;	F:kinase activity;	IEA.																
DR	GO;	GO:0016772;	F:transferase activity;	IEA.																
DR	GO;	GO:0000103;	P:sulfate assimilation;	IEA.																
DR	InterPro;	IPR002891;	APS_kinase.																	
DR	Pfam;	PF01583;	APS_kinase; 1.																	
DR	ProDom;	PD002350;	APS_kinase; 1.																	
DR	TIGRFAMs;	TIGR00455;	_apsK; 1.																	
KW	Kinase.																			
SQ	SEQUENCE 345 AA; 37748 MW; BE4235F256676698 CRC64;																			
Query Match																				
Best Local Similarity 53.1%; Score 873.5; DB 2; Length 345;																				
Matches 188; Conservative 33; Mismatches 68; Indels 65; Gaps 7;																				
QY	10	EPLVTHTQPPSPAP----	GPASOG-----	ORQNTLLSPPTTAVILVNP----	ORAPP 56															
Db	33	DPRATARVVPAAAPVRSRSPANLGLP	PHPRRLRLAPPRITAAVTGGPRRRRAPP 92																	
QY	57	VLPLGLTSDAPLPALVIHGLTPRSSH	SAGLASDGR-----	EGEGRGARTHCHRGIG 110																
Db	93	PL-----	ECAGGSSSLRRPREEEEEESSTAHAGV- 127																	
QY	111	RWRRRRRNGAAGPGEAPHSPVKEKPV	M-----	SNIGKSTNILWHNCLIGQSDQKLLGQKGC 167																
Db	128	-----	SLVGENKVLQMSIVPKASNI	FHDCNAGVQADQKLLGQKGC 169																
QY	168	VVWITGLSGSGKSTLACALSRELHCR	GHLTYVLDGDNLRHGLNRDLS	PKAEDRAENIRRV 227																
Db	170	VVWITGLSGSGKSTLACTDLRELHTR	GKLSVYLDGDNLRHGLNKDLG	FKAEDRAENIRRV 229																
QY	228	GEVAKLFADAGVICIASLISPYRRDR	ACRALLPHSNFIEVFDLPLKICE	ARDPKGLYK 287																
Db	230	GEVAKLFADAGLVCIASFISPYRRDR	SCALLSDGGSFIEVFLNMP	LCESDPKGLYK 289																

```

OC lamids; Gentianales; Apocynaceae; Rauvolfioideae; Vinceae;
OC Catharanthus.
OX NCBI_TaxID=4058;
RN [1]
RN SEQUENCE FROM N.A.
RA Schiffmann S., Schwenn J.-D.;
RT "Isolation of cDNA clones encoding adenosine-5'-phosphosulfate-kinase
RT (SC2.7.1.25) from Catharanthus roseus and an isoform (akn2) from
RT Arabidopsis.";
RL (er) Plant Gene Register PGR98-116.
CC -i- FUNCTION: Catalyzes the synthesis of activated sulfate.
CC -i- CATALYTIC ACTIVITY: ATP + adenylyl sulfate = ADP + 3'-
CC phosphoadenylyl sulfate.
CC -i- PATHWAY: Sulfate activation; cysteine biosynthesis reductive
CC branch; second step.
CC -i- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC -i- SIMILARITY: Belongs to the APS kinase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AF044285; AAC31145.1; -.
CC PIR: T08076; T08076.
CC HSP: Q12657; I07G.
CC InterPro: IPR002891; APS_kinase.
CC Pfam: PF01583; APS_kinase; 1.
CC ProDom: PD002350; APS_kinase; 1.
CC TIGRfams: TIGR00455; apsk; 1.
CC ATP-binding; Chloroplast; Cysteine biosynthesis; Kinase;
KW phosphorylation; transferase; Transient peptide.
KW TRANSIT 1 ? Chloroplast (Potential).
FT CHAIN ? 312 Adenylyl-sulfate kinase.
FT NP BIND 142 149 ATP (Potential).
FT FT ACT_SITE 216 216 Phosphoserine intermediate (By
FT similarity).
SQ SEQUENCE 312 AA; 33656 MW; 16BBD11FB4B1FE27 CRC64;
-----
Query Match 45.6%; Score 833; DB 1; Length 312;
Best Local Similarity 72.9%; Pred. No. 2.3e-53;
Matches 156; Conservative 24; Mismatches 34; Indels 0; Gaps 0;

Qy 130 PVTEKPMVSNIGKSTWILWHNCLIGSDRQKLQKGVVWITGLSGSGKSTLACALSRE 189
Db 99 PGKKILQTTVGNSTWILWHKCAVEKSEQEPQQRGCVIWTGLSGSGKSTLACALSRG 158
Qy 190 LHCRGHLYTVLGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLSPY 249
Db 159 LHAQKLYTILGDNVHGLNSDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLSPY 218
Qy 250 RRDRDRAALLPHSNFIEVFIDLPKICEARDPKGLYKLARTKIKGFTGIDDPYEPFIN 309
Db 219 RKPPDACSRLPEGDFIEVFMVPLKVEARDPKGLYKLARAGKIKGFTGIDDPYEPPLK 278
Qy 310 GEIVIKMKDECPSPKMAKQVLCYLENGYLQA 343
Db 279 SEIVLHQKGMWCDSPCDLADIVISYLENGYLKA 312

RESULT 4
KAP1_ARATH
ID KAP1_ARATH STANDARD; PRT; 276 AA.
AC Q43295;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Adenylyl-sulfate kinase 1, chloroplast precursor (EC 2.7.1.25) (APS
DE kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-
DE phosphosulfate 3'-phosphotransferase).

```

GN Name=AKN1; OrderedLocusNames=At2g14750; ORFNames=F26C24.11, T6B13.1;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 (1)
 RN SEQUENCE FROM N.A.
 RP TISSUE=Leaf;
 RC MEDLINE=94325358; PubMed=80492272; DOI=10.1016/0167-4781(94)90203-8;
 RX Arz H.E., Gisselmann G., Schiffmann S., Schwenn J.-D.;
 RA "A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis
 RT thaliana";
 RL Biochim. Biophys. Acta 1218:447-452(1994).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94345022; PubMed=8066145; DOI=10.1104/pp.105.2.771;
 RA Jain A., Leustek T.;
 RT "A cDNA clone for 5'-adenylylphosphosulfate kinase from Arabidopsis
 RT thaliana";
 RL Plant Physiol. 105:771-772(1994).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta;
 RA Leustek T.;
 RT "Analysis of ATP-sulfurylase and APS-kinase genomic DNA sequences in
 RT Arabidopsis thaliana";
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN (4)
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20083487; PubMed=10617197; DOI=10.1038/45471;
 RX Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
 RA Mofat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
 RA Niemman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
 RA Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana";
 RL Nature 402:761-768(1999).
 CC !- FUNCTION: Catalyzes the synthesis of activated sulfate.
 CC !- CATALYTIC ACTIVITY: ATP + adenylyl sulfate = ADP + 3'-
 CC phosphoadenylyl sulfate.
 CC !- PATHWAY: Sulfate activation; cysteine biosynthesis reductive
 CC branch; second step.
 CC !- SUBCELLULAR LOCATION: Chloroplast.
 CC !- SIMILARITY: Belongs to the APS kinase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X75782; CAA53426.1; -
 DR EMBL; U05238; AAC50035.1; -
 DR EMBL; U59759; AAC50034.1; -
 DR EMBL; AC004705; AAC24182.1; -
 DR PIR; S47640; S47640.
 DR HSP; Q12657; IM7G.
 DR InterPro; IPR002891; APS kinase.
 DR Pfam; PF01583; APS kinase; 1.
 DR ProDom; PD002350; APS kinase; 1.
 DR TIGRFAMs; TIGR00455; apsk; 1.
 KW ATP-binding; Chloroplast; Cysteine biosynthesis; Kinase;
 KW Phosphorylation; Transferrase; Transit peptide.
 FT TRANSIT 1 37 Chloroplast (Potential).
 FT CHAIN 38 276 Adenylyl-sulfate kinase 1.

FT NP_BIND 108 115 ATP (potential).
 FT ACT_SITE 182 182 Phosphoserine intermediate (By
 FT similarity).
 SQ SEQUENCE 276 AA; 29787 MW; CB698643AA09D811 CRC64;
 Query Match 44.7%; Score 816; DB 1; Length 276;
 Best Local Similarity 64.4%; Pred. No. 3.6e-52;
 Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;
 QY 100 GARTCHRGIGRWVRRRRNGAAGEAP---HSPVKEKPVMSNIGKSTNHLWHNCLIGQS 156
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
 Db 43 GSQTLSH-----NKGSIPEVKSGINGHTGQKQGP-LSTVGNSTNWKHECSVEKV 91
 QY 157 DRQKLLGKGCVVWITGLSGSGKSTLACALSELCHLTYVLGDNDLNRHGLNRLDLSFK 216
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
 Db 92 DRQRLDQKGCVVWITGLSGSGKSTLACALNQLYQKGLCYLGDNDVRRHGLNRLDLSFK 151
 QY 217 AEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDCAALLPHSNFIEVFDLPLKI 276
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
 Db 152 AEDRAENIRRVGEVAKLFADAGIICIASLISPYRTDRDACSLLPGEQFVEFMDVPLSV 211
 QY 277 CEARDPKGLYKLARTGKIKGFTGIDDDPYEPPINGEIVIMKDEECPSKAMAKQVLCYLE 336
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
 Db 212 CEARDPKGLYKLARAGIKGFTGIDDDPYEPLNCE--ISLREGGTSPIEMAEKVVGILD 269
 QY 337 ENGYLQA 343
 DB |||||
 Db 270 NKGYLQA 276
 RESULT 5
 Q8LES2 PRELIMINARY; PRT; 276 AA.
 ID Q8LES2
 AC Q8LES2
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative adenosine phosphosulfate kinase.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22089475; PubMed=12093376;
 RA Haas B.J., Volfovsky N., Town C.D., Troughan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation";
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Brover V., Troughan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY085264; AAM62496.1; -
 DR HSP; Q12657; IM7G.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016301; P:kinase activity; IEA.
 DR GO; GO:0016772; P:transferase activity, transferring phosphor. .; IEA.
 DR GO; GO:000103; P:sulfate assimilation; IEA.
 DR InterPro; IPR002891; APS_kinase.
 DR Pfam; PF01583; APS_kinase; 1.
 DR ProDom; PD002350; APS kinase; 1.
 DR TIGRFAMs; TIGR00455; apsk; 1.
 KW Kinase.
 SQ SEQUENCE 276 AA; 29769 MW; 66DF5D2F5C86B0E5 CRC64;
 Query Match 44.7%; Score 816; DB 2; Length 276;
 Best Local Similarity 64.4%; Pred. No. 3.6e-52;
 Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;

RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M.A., Martienssen R., McCombie W.R.,
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana";
 RL Nature 402:769-777(1999).
 CC -!- FUNCTION: Catalyzes the synthesis of activated sulfate.
 CC -!- CATALYTIC ACTIVITY: ATP + adenylyl sulfate = ADP + 3'-
 CC phosphoadenylyl sulfate.
 CC -!- PATHWAY: Sulfate activation; cysteine biosynthesis reductive
 CC branch; second step.
 CC -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).
 CC -!- SIMILARITY: Belongs to the APS kinase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF043351; AAC39520.1; -;
 DR EMBL; AL035708; CAB38907.1; -;
 DR EMBL; AL161596; CAB80657.1; -;
 DR PIR; T06100; T06100.
 DR HSSP; Q12657; 1M7G.
 DR InterPro; IPR002891; APS_kinase.
 DR Pfam; PF01583; APS_kinase; 1.
 DR ProDom; PD002350; APS_kinase; 1.
 DR TIGRPFAM; TIGR00435; apsk; 1.
 DR ATP-binding; Chloroplast; Cysteine biosynthesis; Kinase;
 KW Phosphorylation; Transfrase; Transit peptide.
 FT TRANSIT ?
 FT CHAIN ? 293 Chloroplast (Potential).
 FT NP BIND 142 149 ATP (Potential).
 FT ACT_SITE 136 136 Phosphoserine intermediate (By
 FT similarity).
 SQ SEQUENCE 293 AA; 31977 MW; 71FAD6E9B026886 CRC64;
 Query Match 42.6%; Score 777.5; DB 1; Length 293;
 Best Local Similarity 71.8%; Pred. No. 2.7e-49;
 Matches 145; Conservative 26; Mismatches 30; Indels 1; Gaps 1;
 QY 142 KSTNLHWHNCLIGQSDRQKLLQKGCVMWITGLSGSGKSTLACALSRELHCRGHLTYVLD 201
 Db 91 KAENIVHNESSICRCDRQQLQKGCVMWITGLSGSGKSTVACALSRELHCRGHLTYVLD 150
 QY 202 GNLRLHGLNRLDSFKAEDRAENIRRVGEVAKLFADAGVICTASLISPYRRDRDRAALLP 261
 Db 151 GDNVHGLNRLDTFKAERHTENIRRVGEVAKLFADGVICTASLISPYRRDRDRAALLP 210
 QY 262 HSNFTLVFDLPLKICEARDPKGLYKLTARKTGKGTGIDDPYEPPIGEIVK-WKDEE 320
 Db 211 DGFDFEVMDFVPLHVCESSDPPKGLYKLTARKTGKGTGIDDPYEPPIGEIVK-KHTGDDE 270
 QY 321 CPSKAMAKQVLCYLEENGVLQ 342
 Db 271 SCSPRQMAENIISYLNQKYLE 292
 RESULT 8
 Q84JF0
 ID Q84JF0 PRELIMINARY; PRT; 310 AA.
 AC Q84JF0
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Putative adenylylsulfate kinase.

GN Name=At5g67520;
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
 RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
 RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
 RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
 RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
 RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
 RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
 RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
 RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BT0031977; AAO42019.1; -;
 DR EMBL; BT005193; AAO50726.1; -;
 DR HSSP; Q12657; 1M7G.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0016772; F:transferase activity, transferring phosphor. .; IEA.
 DR GO; GO:000103; P:sulfate assimilation; IEA.
 DR InterPro; IPR002891; APS_kinase.
 DR Pfam; PF01583; APS_kinase; 1.
 DR ProDom; PD002350; APS_kinase; 1.
 DR TIGRPFAMs; TIGR00455; apsk; 1.
 KW Kinase.
 SQ SEQUENCE 310 AA; 34064 MW; CIAC01A4242A0BDF CRC64;
 Query Match 41.5%; Score 757.5; DB 2; Length 310;
 Best Local Similarity 66.4%; Pred. No. 8.6e-48;
 Matches 144; Conservative 29; Mismatches 29; Indels 15; Gaps 2;
 QY 141 GKSTNLHWHNCLIGQSDRQKLLQKGCVMWITGLSGSGKSTLACALSRELHCRGHLTYVLD 200
 Db 84 GKQKNVHWHDCPVTKSDRQELIKQKGCVMWITGLSGSGKSTLACALSRELHCRGHLTYVLD 143
 QY 201 DGDNLHGLNRLDSFKAEDRAENIRRVGEVAKLFADAGVICTASLISPYRRDRDRAALLP 260
 Db 144 DGDNLHGLNRLDSFKAEDRAENIRRVGEVAKLFADAGVICTASLISPYRRDRDRAALLP 203
 QY 261 PHSNFTLVFDLPLKICEARDPKGLYKLTARKTGKGTGIDDPYEPPIGEIVK----- 315
 Db 204 PQGDFIEVDFVPLHVCESSDPPKGLYKLTARKTGKGTGIDDPYEPPIGEIVKQNSRDK 263
 QY 316 -----MKDEECPSPKA-----MAKQVLCYLEENGVLQ 342
 Db 264 GLSSSSSSSSSSSSSSSSSSSSSLCEMADIVWSYLDQNGYLK 300
 RESULT 9
 Q9FJX1
 ID Q9FJX1 PRELIMINARY; PRT; 290 AA.
 AC Q9FJX1
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Adenylylsulfate kinase-like protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsie.

```

OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:203-216(1998).
DR EMBL; AB013390; BAB08460.1; -.
DR HSSP; Q12657; 1M7G.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016772; F:transferase activity, transferring phosphor. .; IEA.
DR GO; GO:0000103; P:sulfate assimilation; IEA.
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_kinase; 1.
DR ProDom; PD002350; APS_kinase; 1.
DR TIGRFAMs; TIGR00455; apsK; 1.
KW Kinase.
SQ SEQUENCE 290 AA; 32195 MW; BB7FB12990A6F6F8 CRC64;

Query Match 41.3%; Score 755; DB 2; Length 290;
Best Local Similarity 69.8%; Pred. No. 1.2e-47;
Matches 141; Conservative 28; Mismatches 33; Indels 0; Gaps 0;

QY 141 GKSTNHLWNCLIGSDRQKLLGQKGVVWITGLSGSKSTLACSLRELHCRGHLTYVL 200
DB 79 GKQKNIVVHDCPVTKSDRQELIKHGCVIWTGLSGSKSLACSLRALHNRGKLSYIL 138

QY 201 DGDNLRLHGLNRLDLSFKAEADRAENIRRVGEVAKLFADAGVICIASLISPYRRDRACRALL 260
DB 139 DGDNVRHGLNSDLSFEADRAENIRRVGEVAKLFADSGIICIASLISPYRTERAACRALL 198

QY 261 PHSNFIIEVDLPLKICEARDPKGLYKLRAGKIKGFTGDDPYEPPINGEIVIKMDDEE 320
DB 199 PQGDFIEVFMDVPLHVCARDPKGLYKRRAGKIKGFTGDDPYEAPLDCVHIISNFS 258

QY 321 CPSPKAMAKVLCYLEENGYLQ 342
DB 259 SSSLCEMADIVVSYLDQNGYLK 280

RESULT 10
Q8LF64 PRELIMINARY; PRT; 305 AA.
AC Q8LF64;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Adenylsulfate kinase-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsi.
NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22088475; PubMed=12093376;
RA Haas B.-J., Volfsovs N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085031; RAM61589.1; -.
DR HSSP; Q12657; 1M7G.
DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016772; F:transferase activity, transferring phosphor. .; IEA.
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_kinase; 1.
DR ProDom; PD002350; APS_kinase; 1.
DR TIGRFAMs; TIGR00455; apsK; 1.
FT NON TER 152 152
SQ SEQUENCE 152 AA; 16974 MW; DBAC812B71D50DES CRC64;

Query Match 34.6%; Score 632; DB 2; Length 152;
Best Local Similarity 77.0%; Pred. No. 6.5e-39;
Matches 117; Conservative 20; Mismatches 15; Indels 0; Gaps 0;

QY 137 MSNIGKSTNHLWNCLIGSDRQKLLGQKGVVWITGLSGSKSTLACSLRELHCRGHL 196
DB 1 MSTVGNSTNIFWQSPIGKTERQKLLNQKGVVWITGLSGSKSTLACSLRELNRGKL 60
```

[illegible]

```

316 MKDECPSPKAMAKQVLCYLEENGVL 341
      :|:|
      ::|:|::|
181 DTEBKC-----VDNIIEYLKQHKIL 200

RESULT 13
Q9NDP8
ID Q9NDP8 PRELIMINARY; PRT; 618 AA.
AC Q9NDP8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ATP sulfurylase/APS kinase.
GN Name=Ci-ASAK;
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Cionidae; Ciona.
NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RA Hotta K., Takahashi H., Asakura T., Saitoh B., Takatori N., Satou Y.,
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036852; BAB00629.1; -.
DR HSSP; Q12657; 1M7G
DR GO; GO:0005622; C:intracellular; ISS.
DR GO; GO:0005524; F:ATP binding; ISS.
DR GO; GO:0016779; F:nucleotidyltransferase activity; ISS.
DR GO; GO:0004781; F:sulfate adenylyltransferase (ATP) activity; ISS.
DR GO; GO:0016740; F:transferase activity; ISS.
DR GO; GO:0001501; P:skeletal development; ISS.
DR InterPro; IPR002891; APS_kinase.
DR InterPro; IPR002650; ATP-sulfurylase.
DR Pfam; PF01583; APS_kinase; 1.
DR Pfam; PF01747; ATP-sulfurylase; 1.
DR TIGRFAMs; TIGR00455; apsk; 1.
KW Kinase.
SQ SEQUENCE 618 AA; 69559 MW; 75223CCB022BC297 CRC64;

```

Matches	114;	Conservative	30;	Mismatches	51;	Indels	14;	Gaps
QY	122	APGEAPH-SPVKEKPYMSNTGKSNILWHNCLIGQSDRQKLLQ---	KGCVVWITGLSGS	177				
Db	2	ASATANHGDIIQEKPPM-----SSNIYVQQHVSRRDERGQVMGKGA	FGCTVWFTGLSGA	56				
QY	178	GKSTLACALSGRELHCRGHLETVYLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADA	237					
Db	57	GKSTLSMKLEYLCSKGIPIAYSLDGDNIRHGLNKDLGFAPADRENIIRI	GEVAKLFADA	116				
QY	238	GVICIASLISPYRRDRDCA	RALLPHSN--FTEVFIDLPKTC	EARDPRGLYKLA	RTGKIK	295		
Db	117	GVVCLVSFISPYRKORQSARRVHNKSNLP	FTIEVFVDTPLOVCEGRDVKGLYKKA	RQGLK	176			
QY	296	GFTGIDDDPYPPINGEIVIKMKD---	BEC	321				
Db	177	GFTGIDSPYPPENAEVVTKTADMP	VEDC	205				

```
Db          177 GFtGIDSPYEPPENAEVVTKTADMPVEDC 205
                                     :
RESULT 14
CYC2 BACSU
ID _CYC2_BACSU      STANDARD;      PRT;   199 AA.
AC 006735;
```

30-MAY-2000	(Rel. 39, Created)
30-MAY-2000	(Rel. 39, Last sequence update)
25-OCT-2004	(Rel. 45, Last annotation update)
DE	Probable adenylyl-sulfate kinase (EC 2.7.1.25) (APS kinase)
DE	(Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate
DE	3'-phosphotransferase).
GN	Name=yisZ; OrderedLocusNames=BSU10910;
OS	Bacillus subtilis.
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxID=1423;	
11]	
SEQUENCE FROM N.A.	
RC	STRAIN=168;
RC	MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RC	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RC	Azevedo V., Berto M.G., Bessieres P., Bolotin A., Borchert S.,
RA	Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA	Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA	Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA	Daniot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA	Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA	Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,
RA	Ghim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
RA	Guiseppi G., Guy B.J., Haga K., Hatched J., Harwood C.R., Henaut A.,
RA	Hilbert H., Holgappel S., Hosono S., Hulio M.F., Itaya M.,
RA	Jones L.-M., Joris B., Karamata D., Kasahara Y., Kleerr-Blanchard M.,
RA	Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA	Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA	Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA	Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA	Noback M., Needing D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA	Park S.H., Parro V., Pohl T.M., Portetalle D., Porwollik S.,
RA	Prescott A.M., Presecan E., Fujic P., Furnelle B., Rapoport G.,
RA	Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA	Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA	Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P.,
RA	Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA	Takenaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P.,
RA	Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA	Vassarotti A., Viari A., Wambutt R., Wedler H., Wedler H.,
RA	Weitznegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA	Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RA	Yoshikawa H., Danchin A.,
RT	"The complete genome sequence of the Gram-positive bacterium Bacillus
RT	subtilis.";
RL	Nature 390:249-256(1997).
CC	-I- FUNCTION: Catalyzes the synthesis of activated sulfate.
CC	-I- CATALYTIC ACTIVITY: ATP + adenylyl sulfate = ADP + 3'-
CC	phosphoadenylyl sulfate.
CC	-I- PATHWAY: Sulfate activation; cysteine biosynthesis reductive
CC	branch; second step.
CC	-I- SIMILARITY: Belongs to the APS kinase family.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/)
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; Y09476; CNA70655.1; ALT_INIT.
DR	EMBL; Z99109; CAB12931.1; -.
DR	PIR; A69839; A69839.
DR	HSSP; O12657; 1M7G.

Db ||||| ||||| | : | : ||| | : ||| | ||||| | | : ||||| |||||
122 ISPYVRDRDAIRATLDDGDFQBFIFDTPIEICEKRDPKGLYKKARAGEIKGFTGIDDPYE 181
Qy ||||| ||||| | : | : ||| | : ||| | ||||| | | : ||||| |||||
306 PPINGEIVIKMKDEECPCPKAMAKQVLCYLEENGYL 341
Db ||||| ||||| | : | : ||| | : ||| | ||||| | | : ||||| |||||
182 APPRPELRL---DGAAKDAETLAEVIAHLERVGVI 214

Search completed: September 9, 2005, 13:03:03
Job time : 122 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 9, 2005, 13:07:17 ; Search time 665 Seconds
(without alignments)
3387.314 Million cell updates/sec

Title: US-10-829-432-4

Perfect score: 1826

Sequence: 1 RPFHINQTEPLVTHQTPP.....PKMAKQVLCYLENGYLQA 343

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10829432/runat_08092005_132224_25132/app_query.fasta.1.519
-DB=Published Applications NA -QMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10829432.cgn_1_1.723 @runat_08092005_132224_25132
-NCPUS=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBIOLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1826	100.0	1217	19	US-10-829-432-3	Sequence 3, Appli
2	1660.5	90.9	1291	20	US-10-425-115-116953	Sequence 116953,
3	1415	77.5	1335	18	US-10-425-114-941	Sequence 941, App
4	1250.5	68.5	1498	20	US-10-425-115-116950	Sequence 116950,
5	1008.5	55.2	1319	20	US-10-739-930-5324	Sequence 5324, Ap
6	944	51.7	945	19	US-10-437-963-51779	Sequence 51779, A
7	936	51.3	945	19	US-10-767-701-9302	Sequence 9302, Ap
8	905.5	49.6	1028	18	US-10-425-114-16540	Sequence 16540, A
9	896.5	49.1	2611	20	US-10-425-115-62889	Sequence 62889, A
10	892.5	48.9	1180	18	US-10-425-114-18247	Sequence 18247, A
11	887	48.6	890	19	US-10-829-432-1	Sequence 1, Appli
12	876	48.0	928	19	US-10-829-432-9	Sequence 9, Appli
13	875.5	47.9	1312	18	US-10-425-114-1150	Sequence 1150, Ap
14	854.5	46.8	1764	20	US-10-425-115-135127	Sequence 135127,
15	854	46.8	1116	18	US-10-425-114-9652	Sequence 9652, Ap
16	854	46.8	1256	18	US-10-424-599-62905	Sequence 62905, A
17	853	46.7	981	18	US-10-424-599-119620	Sequence 119620,
18	852	46.7	936	19	US-10-829-432-7	Sequence 7, Appli
19	846.5	46.4	1302	18	US-10-425-114-27400	Sequence 27400, A
20	816	44.7	831	9	US-09-938-842A-710	Sequence 710, App
21	816	44.7	831	11	US-09-938-842A-710	Sequence 710, App
22	788	43.2	687	17	US-10-260-238-869	Sequence 869, App
23	783.5	42.9	1821	20	US-10-425-115-35036	Sequence 35036, A
24	778.5	42.6	920	16	US-10-342-224-65	Sequence 65, Appl
25	770.5	42.2	548	19	US-10-437-963-72904	Sequence 72904, A
26	690	37.8	1132	18	US-10-424-599-119621	Sequence 119621,
27	640	35.0	378	17	US-10-260-238-5283	Sequence 5283, Ap
28	615.5	33.7	635	20	US-10-425-115-35035	Sequence 35035, A
29	571	31.3	603	17	US-10-282-122A-15460	Sequence 15460, A
30	561	30.7	595	17	US-10-260-238-4843	Sequence 4843, Ap
31	555	30.4	600	17	US-10-369-493-41371	Sequence 41371, A
32	534.5	29.3	609	17	US-10-369-493-40905	Sequence 40905, A
33	530	29.0	1592	19	US-10-437-963-45866	Sequence 45866, A
34	521.5	28.6	597	17	US-10-282-122A-34804	Sequence 34804, A
35	521.5	28.6	615	22	US-10-724-972A-1295	Sequence 1295, A
36	519.5	28.5	588	17	US-10-369-493-40371	Sequence 40371, A
37	516	28.3	1863	17	US-10-369-493-34422	Sequence 34422, A
38	514.5	28.2	591	17	US-10-282-122A-10327	Sequence 10327, A
39	514.5	28.2	612	17	US-10-282-122A-32765	Sequence 32765, A
40	510.5	28.0	1842	17	US-10-369-493-39203	Sequence 39203, A
41	510.5	28.0	1869	17	US-10-369-493-39946	Sequence 39946, A
42	510.5	28.0	1890	17	US-10-369-493-39575	Sequence 39575, A
43	510	27.9	642	17	US-10-282-122A-41542	Sequence 41542, A
44	505.5	27.7	597	17	US-10-282-122A-35733	Sequence 35733, A
45	503	27.5	603	17	US-10-282-122A-22900	Sequence 22900, A

ALIGNMENTS

RESULT 1
US-10-829-432-3
; Sequence 3, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14

2


```
QY 135 oValMetSerAsnIleGlyLeuSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyG1 155
DB 345 TGTAAATGTCACAAATTTGGGAAATCGACTAATATTTATGGCAATTTGCTGATTGGACA 404
QY 155 nSerAspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSe 175
DB 405 ATCTGATAGACAGAAATTCCTGGGCAAAAAGGCTGTGTGATGATGATGATGATGATGATG 464
QY 175 rGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHi 195
DB 465 TGGTTCAGGGAAGAAGTACTTCTGATGTCACCTGAGTCGTGAGTGTGATGTCATGTCAGAGGCCA 524
QY 195 sLeuThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPh 215
DB 525 CCTCACGTATGTACTTGTGATGTTGACAACTCAGACATGCGCTAAATAGAGATTTAAGCTT 584
QY 215 eLysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAl 235
DB 585 TAAGGAGAGAACCGCTGCAGAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTC 644
QY 235 aAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAspAl 255
DB 645 TGATGCTGTGTGATGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 704
QY 255 aCysArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLy 275
DB 705 ATGCGTGTCTACTTCCACATTTCACTTTATGAAGTATTTATGATTTGCCCTAAA 764
QY 275 sLysCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLy 295
DB 765 AATTTGTGAAGCTCGTGATCTCTAAAGGCTTATACAGCTTGACAGCTTACAGGAAGATTAA 824
QY 295 sGlyPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValIleLy 315
DB 825 AGTTTCTACTGGAATTTGATGATCCATACCAACCAACCAATTAATGATGATGATGATGATG 884
QY 315 sMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrIle 335
DB 885 GATGAAGATGGGAATGCCCTTACCCAAAGCAATGGCCAGCAAGTTCATGCTACTCT 944
QY 335 uGluGluAsnGlyTyrLeuGlnAla 343
DB 945 TGAAGAGAACGGATATTTGCAAGCT 969
```

RESULT 3

```
US-10-425-114-941
; Sequence 941, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 941
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700097023_FLI
US-10-425-114-941
```

Alignment Scores:

```
Pred. No.: 5,76e-128 Length: 1335
Score: 1415.00 Matches: 278
```

```
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 89.68% Mismatches: 3
Query Match: 77.49% Indels: 28
DB: 18 Gaps: 1
US-10-829-432-4 (1-343) x US-10-425-114-941 (1-1335)
QY 62 ThrProSerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrProArgSerSer 81
DB 2 ACCCTAGCGATGGCCACTCCCGCGCTCGTGATCCATGGCTCACTCCCGCTTCTCTCA 61
QY 82 HisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgGluGlyGluGlyArgGlyAla 101
DB 62 CACTCTTCGGCGGCTCGCCAGCATAGTGGCGCGCGAGGGGAGGCGCGCGTGG 121
QY 102 ArgThrHisCysHisAsnGlyIleGlyValArgTrpValArgArgArgArgAsnGlyAla 121
DB 122 CGTACGCACTGCCACCGCGCATTTGGCGGTAGGTGCGCGCGCGCGCGGGAATGAGCA 181
QY 122 AlaPro-GlyGluAlaProHisSerProValLysGlyLysProValMetSerAsnIleG1 141
DB 182 GCGCCACGCGGAGGCGCCGACAGCCAGTGAAGGAGAAGCCTGTAATGTCAACATTGG 241
QY 141 yLysSerThrAsnIleLeuTrpHisAsnCysIleGlyGlnSerAspArgGlnLysIle 161
DB 242 GAAATCGACTAATATTTTATGGCAATTTGCTTGTATGGACAATCTGATAGACAGAAAT 301
QY 161 uLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerTh 181
DB 302 GCTGGGCAAAAAGCTGTGCTGATGATTAACAGACTTCAGTGGTTCAGGGAAGAATAC 361
QY 181 rLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAs 201
DB 362 TCTTGATGTGCACTGAGTGTGAGTTCATTGCAGAGAGCCACCTCACGTATGACTTGA 421
QY 201 pGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAl 221
DB 422 TGGTGACAACTCAGACATGGCTTAAATAGAGATTTAAGCTTTAAGGAGAGACCGTGC 481
QY 221 aGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCy 241
DB 482 AGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTCTGATGCTGGTGTCTATATG 541
QY 241 sIleAlaSerLeuIleSerProTyrArgArgAspArgAlaCysArgAlaLeuLeuPr 261
DB 542 CATTTGCTAGCTTGATATCTCCATACAGAGAGATGCTGATGATGCTGCTTACTTCC 601
QY 261 oHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAs 281
DB 602 ACATTTCTAACTTTATTGAAGTATTTATTGATTTGCCCTTAAATAATTTGTGAAGCTCGTGA 661
QY 281 pProLysGlyLeuTyrLysLeuAlaAsnThrGlyLysIleLysGlyPheThrGlyLys 301
DB 662 TCCTTAAGGCGCTGTACAAGCTTGACAGCTTGACAGGAAGATTAAGGTTTCACTGGAATGGA 721
QY 301 pAspProTyrGluProProIleAsnGlyGlu----- 311
DB 722 TGATCCATACGACCAACCAATTTAATGTTGAGATAAGGCAATAGTTATATACATCTGGCAT 781
QY 312 -----lleValIleL 315
DB 782 TGTTTTCAAGGATTAACCTTTGAACTCAGAACGCTCTTCTCTGCGAGTTTAAAGATAGTAATTA 841
QY 315 yMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrL 335
DB 842 CGATGAAGAAGTGGGAATTCCTTCCACCAAGCAATGGCCAAAGCAATGGCCAAAGTTCCTATGCTACC 901
QY 335 euGluGluAsnGlyTyrLeuGlnAla 343
DB 902 TTGAAGAGAACGGATATTTTGAAGCT 927
RESULT 4
US-10-425-115-116950
```

```
; Sequence 116950, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 116950
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_38149C.1
US-10-425-115-116950

Alignment Scores:
Pred. No.: 7,18e-112 Length: 1498
Score: 1250.50 Matches: 263
Percent Similarity: 78.70% Conservatative: 3
Best Local Similarity: 77.81% Mismatches: 11
Query Match: 68.48% Indels: 64
DB: 20 Gaps: 3

US-10-829-432-4 (1-343) x US-10-425-115-116950 (1-1498)

QY 9 ThrGluProLeuValThrHisThrGlnGlnProProSerProAlaProGlyProAlaSer 28
DB 22 ACAGAACCTCTGTCA--CACAGCAGCAAGCAGCCAGCCAGCCGCCGCCACCCAC- 78
QY 29 GlnGlyGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 48
DB 79 -----GGCAACACCTCTCTCAGCCGAGCCGAGCCGAGCCGAGCCGAGCCG 123
QY 49 ValAsnProGlnArgAlaProValLeu-----ProGlyLeuThrProSerAspAla 66
DB 124 GAATGTCACAGCGCGCGCTCCCGTCTCCAGTCCAGTCCAGTCTCCAGTCCAGTCC 183
QY 67 ProLeuProAlaLeuValHisGlyLeuThrProArgSerSerHisSerSerAlaGly 86
DB 184 CCACCTCCCGCGCTCGTGATCATCGGCTCATCTCCCGTCTCTCACACTTCCGCGGC 243
QY 87 LeuAlaSerAspSerGlyArgArgGluGlyGluGlyGluGlyGluGlyGluGlyGly 106
DB 244 CTCGCCAGCGATAGTGGCGCGCGCGAGGGGAGGGCGCGGTCGCGTACGCACTGCC 303
QY 107 ArgGlyIleGlyArgTrpValArgArgArgArgArgArgArgArgArgArgArg 126
DB 304 CCGCGCATTTGGCGGTAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 363
QY 126 aProHisSerProValHisGlyLeuProValMetSerAsnIleGlyLeuSerThrAsn 146
DB 364 CCGCACAGCCAGTGAAGAGAGAGCTGTATGTCAAACTTGGAAATCGACTAATAT 423
QY 146 eLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuGlyGln 166
DB 424 TTTATGGCAAAATTTGTTGGAATCTCTATAGACAGAAATTTGCTGGGCAAAAAAG 483
QY 166 vCysValValTrpIleThrGlyLeuSerGlySerGlyLeuSerThrLeuAlaCysAla 186
DB 484 CTGTGTCGTATGATTAACAGGACTCAGTGTTCAGGAAAAGTACTCTTTCATGTG 543
QY 186 uSerArgGluLeuHisCysArgGlyHisLeuThrTrpValLeuAspGlyAspAsnLeu 206
DB 544 GAGTCGTGATTTGATTCAGAGGCCACCTCAGTATGTACTTGTATGGTGAACCTC 603
QY 206 gHisGlyLeuAsnArgAspLeuSerPheIysAlaGluAspArgAlaGluAsnIleArg 226
DB -----
```

```
Db 604 ACATGGCCCTAAATAGAGATTTTAAGCGCAGAAAGACCGTCGAGAAAAATATACGAAG 663
QY 226 qValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuI 246
Db 664 AGTTGGTGAAGTGGCAAGCTTTTCTGATGCTGGTGTCTATATGCAATGCTAGCTT 723
QY 246 eSerProTrpArgArgAspArgAlaCysArgAlaLeuLeuProHisSerAsnPheI 266
Db 724 ATTGACATACAGGAGAGA----- 741
QY 266 eGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeu 286
Db 741 ----- 741
QY 286 rLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspProTyrGlu 306
Db 742 -----CC 743
QY 306 oProIleAsnGlyGluIleValIleLysMetLysAspGluCysProSerProLysAl 326
Db 744 ATCAATTAATGTTGAGATAGTAATTAAGATCTAAGATGGGAATGCCCTTCCACATA 803
QY 326 aMetAlaLysGlnValLeuCysTyrLeuGluGlnAsnGlyTyrLeuGlnAla 343
Db 804 AATGCCCAAGCAAGTTCTATGCTACCTTGAAGAGAACGATATTTGCAAGCT 855

RESULT 5
US-10-739-930-5324
; Sequence 5324, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 5324
; LENGTH: 1319
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER7564_1
US-10-739-930-5324

Alignment Scores:
Pred. No.: 2,4e-88 Length: 1319
Score: 1008.50 Matches: 203
Percent Similarity: 80.80% Conservatative: 20
Best Local Similarity: 73.55% Mismatches: 40
Query Match: 55.23% Indels: 13
DB: 20 Gaps: 4

US-10-829-432-4 (1-343) x US-10-739-930-5324 (1-1319)

QY 73 lIleHisGlyLeuThrProArgSerSerHisSerSerAlaGlyLeuAlaSerAspSer 92
Db 4 ATTCAC---GTCACGCGCGCACATTTTATTATTAGTCTGCGACGAGCAGTGGCTGT 60
QY 93 ArgArgGluGlyGluGlyArgGlyAlaArgThrHisCysHisArgGlyIleGlyArg 112
Db 61 CGAGTCGAGCAGCTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 99
QY 113 ValArgArgArgArgArgAsnGlyAlaAlaProGlyGluAlaProHisSerPro 130
Db 100 GAGCTTCCTCCCTCCACTCCGTCAGCTCCCGGGCGCTCCCGCGCGCGCGCGCG 159
QY 131 -----ValLysGluLysProValMetSerAsnIleGlyLysSerThrAsnIle 147
Db 160 CGCGCGCGAGTGCAGAGAGAGCGCTGTCATGTCAAACTTGGGAAATCACTAAT 219
QY 148 TrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuGlyGlnLysGly 167
Db -----
```

Db 220 TGGCATGCTGCCCAATGGACAACTGACGACAGAAATTCGTGGGGCAAAAGGATGT 279
Qy 168 ValValTrrpIleThrClyLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSer 187
Db 280 GTCATATGATCACAGATCTACGCGGTTTACGGGAAAGTACCGTTGCTGCTACCTGAGT 339
Qy 188 ArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGlyAspAsnLeuArgHis 207
Db 340 CGGGAATTACCTACAGAGGCCACACACGATATGCTTGTGATGGTGACAACTCAGACAT 399
Qy 208 GlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgVal 227
Db 400 GGCCTCAATCGAGATCTAAGCTTCAAGGCGAGAGACCGTACAGAAAATATACGAAAGTT 459
Qy 228 GlyGluValAlaIysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSer 247
Db 460 GGTGAAGTGGCAAGCTTTTTCAGATGCTGGTACCATATGATTCGTAGTTGATATCT 519
Qy 248 ProTyrArgArgAspArgAlaCysArgAlaLeuLeuProHisSerAsnPheIleGlu 267
Db 520 CCATACAGGAGACCGTATGATGTCGCGAGCTTACTTCCAGATTCATAGATTATGNA 579
Qy 268 ValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLys 287
Db 580 GTATTATGGAATTTGCCACTAGAAATATGTAAGCTCGTGAATCTTAAAGGGTTATACAAG 639
Qy 288 LeuAlaArgThrClyLysIleLysGlyPheThrGlyIleAspAspProTyrGluProPro 307
Db 640 CTTGACCGCACAGGAAGAAATTAAGGGTTTACCGGAGTTGATGATCCATACGAATCACCA 699
Qy 308 IleAsnGlyGluIleValIleLysMetLysAspGluGluCysProSerProLysAlaMet 327
Db 700 GTGATAGTGAGATGATTAATGAGTGAAGTGGGGAATGCCCTTACCGAAGCAATG 759
Qy 328 AlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuGlnAla 343
Db 760 GCCAGCAAGTTCTATCTCTACCTTGAGAGAAATGGATATTTGCAGGCT 807

RESULT 6

US-10-437-963-51779
; Sequence 51779, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 51779
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5413C.1

US-10-437-963-51779

Alignment Scores:
Pred. No.: 3e-82 Length: 915
Score: 944.00 Matches: 194
Percent Similarity: 74.01% Conservatives: 31
Best Local Similarity: 63.82% Mismatches: 59
Query Match: 51.70% Indels: 20
DB: 19 Gaps: 5

US-10-829-432-4 (1-343) x US-10-437-963-51779 (1-915)
Qy 58 LeuProGlyLeuThrProSerAspAlaProLeuProAlaLeuValIleHis----- 74
Db 5 TGGCAGGAGATCTCCCATCCCGTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTG 64
Qy 75 -----GlyLeuThrPro-ArgSerSerHisSerSerAlaGlyLeuAlaSer----- 89
Db 65 CTGCCGATGGCGCTCTCCCGGAGGGGTACTGAGGCCACCGTGGGGAGGAGTAGTGTGCCG 124
Qy 90 -AspSerGlyArgArgGluGlyGluGlyAlaArg-ThrHisCysHisArgGlyI 109
Db 125 GCGTGAGGCTCGTCAGCGCGAGGCAATGAGCGTCTGACGCGCGCGCGCGCGCGGA- 183
Qy 109 leGlyArgTrrpValArgArgArgArgAsnGlyAlaAla---ProGlyGluAlaProH 128
Db 184 --AGCAGCTGGGACAGGAAGGCTGCGCGTCCGAGCGCGCGCGCGCGCGCGCGCG 241
Qy 128 IsSer-----ProValLysGluLysProValMetSerAsnI 140
Db 242 TCGACGGGAAGGACGACGACGATGTGTGCCAGTTTTTAAACNAGCTCATGACTCACTG 301
Qy 140 leGlyLysSerThrAsnIleLeuTrrpHisAsnCysLeuIleGlyGlnSerAspArgGlnL 160
Db 302 TCGGTAAATCAACAAACATCCGATGGCATGCTGCCAGTAAACCCAGCTTGACAGACAAA 361
Qy 160 ysLeuLeuGlyGlnLysGlyCysValValTrrpIleThrGlyLeuSerGlySerGlyLys 180
Db 362 AGTTCTCAACCCAGAAAGGCTGTGTGTGTGGATCCTGCGCTAAGTGGTTCCAGGAAAA 421
Qy 180 erThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrrpVal 200
Db 422 GCACCTTTCATGTGCTGCTGAGCTGAGCTGCACTCAAGAGGGCATCTGACCTATGTTC 481
Qy 200 euAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspA 220
Db 482 TTGACGGCGACAATCTCCGCGACGCGCTGAAACAAAGATCTCAGCTTCAAAGCCAAGATC 541
Qy 220 rgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 240
Db 542 GTGCCGAAAATATACGACAGAGTTGGAGAAAGTGGCAAGCTGTTGCAGATGCTGGATTGA 601
Qy 240 leCysIleAlaSerLeuIleSerProTyrArgArgAspArgAlaCysArgAlaLeuL 260
Db 602 TCTGCATTTACTAGTTTGTATATACCTATAGATGATCGAAGCGCTGCCCAATTAC 661
Qy 260 euProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaA 280
Db 662 TACCAATTCCTCGTTTCATGAGGTGTTCTTGAATGTCCCACTTGAAGTATGTGAAGAAA 721
Qy 280 rgAspProLysGlyLeuTrrpLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyI 300
Db 722 GGGATCCAAAAGGCGCTGACAAAGCTGCTCGCGCGCAAAATCAAGGCTTTTACGGGAA 781
Qy 300 leAspAspProTyrGluProLysLeuGlyGluIleValIleLysMetLysAspGluG 320
Db 782 TAGATGATCCTTATGAAAACACCTTCAGATTGTGAGATTGTATACGATCGCAAGTTGGG 841
Qy 320 luCysProSerProLysAlaMetAlaLysGlnValLeuCysTrrpLeuGluAsnGlyT 340
Db 842 ACTGCCCTTACCTAAATCAATGGCTGATCAAGTAGTGTCTATCTTGAAGCCAATGAT 901
Qy 340 yrLeuGln 342
Db 902 TCTTTCAG 909

RESULT 7

US-10-767-701-9302
; Sequence 9302, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 9302
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS114933_1
US-10-767-701-9302

Alignment Scores:
Pred. No.: 1.88e-81 Length: 945
Score: 936.00 Matches: 178
Percent Similarity: 96.81% Conservativity: 4
Best Local Similarity: 94.68% Mismatches: 6
Query Match: 51.28% Indels: 0
DB: 19 Gaps: 0

US-10-829-432-4 (1-343) x US-10-767-701-9302 (1-945)

```
QY 156 SerAspArgGlnLysLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSer 175
DB 2 TCTGATAGACAGAGTTGCTGGGACAAAAGCTGTGCTGTGATACAGGACTCAGT 61
QY 176 GlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHis 195
DB 62 GGTTCAGGGAAAAGTACTCTTCTGTGCACTGAGTCATGAGTTGCATTATAGAGCCAC 121
QY 196 LeuThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhe 215
DB 122 ATCAGCTATGATCTTGATGGTGACAACTTAGACATGGCCTCAATCGAGATTTAAGCTTT 181
QY 216 LysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAla 235
DB 182 AAGGCAGAAAGCCGTGCAGAAATATACGAAGATTTGTGAAGTGGCAAGCTTTTTCGCC 241
QY 236 AspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgAspArgAspAla 255
DB 242 GATGCTGGTATCATATGCAATTCAGTTGATATCTCCATACAGAGAGATCGTGATGCA 301
QY 256 CysArgAlaLeuLeuProHisSerAsnPhelIleGluValPheIleAspLeuProLeuLys 275
DB 302 TGCCGTGCTCTACTTCCAGATTCTAACTTTATTGAAGTATTTATGGATTTGCCCTAAAA 361
QY 276 IleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLys 295
DB 362 ATTGTGAAGCTCGTGATCCTAAAGCTGTGACAAGCTTGCACGACAGAAAGATTAAA 421
QY 296 GlyPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValIleLys 315
DB 422 GGTTCCTCAGTGAATTTGATGATCCATACGAAACGCCAGTAAATGTTGAGATAGTAACTACG 481
QY 316 MetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeu 335
DB 482 ATGAAGAAGATGGGAATGCCCTTACCCAAAGCAATGGCGAAGCAAGTCTTATCATACCTT 541
QY 336 GluGluAsnGlyTyrLeuGlnAla 343
DB 542 GAAGAGACCGATATTTGCAAGCT 565
```

RESULT 8

US-10-425-114-16540
; Sequence 16540, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Tongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 16540
; LENGTH: 1028
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3066-014-B1_FLI
US-10-425-114-16540

Alignment Scores:
Pred. No.: 1.95e-78 Length: 1028
Score: 905.50 Matches: 187
Percent Similarity: 75.09% Conservativity: 21
Best Local Similarity: 67.51% Mismatches: 57
Query Match: 49.59% Indels: 13
DB: 18 Gaps: 2

US-10-829-432-4 (1-343) x US-10-425-114-16540 (1-1028)

```
QY 71 LeuValIleHisGlyLeuThrProArgSerSerHisSerAlaGlyLeuAlaSerAsp 90
DB 11 CTCCTCCGCCCGCGCTCTCCAGCGCGTTTGCATCGCCCGCACATCAAGAGCGG 70
QY 91 SerGlyArgGluGlyGluGlyArgGlyAlaArgThrHisCysHisArgGlyIleGly 110
DB 71 CGCGCTCGCGCGCAGCGCGGGACGA-----CGAGGACGACA 109
QY 111 ArgTrpValArgArgArgAsnGlyAlaAlaProGlyGluAlaProHis----- 128
DB 110 CGACGGTCGCGCGCGCGCGCAGAG-GCCGCTTCTAAGCGGTTCAGCGCGCGCGCGCTC 168
QY 129 -----SerProValLysGluLysProValMetSerAsnIleGlyLysSerThrAsn 145
DB 169 CGAGGGATCAGCAGCAGCAGCAGCGCGCTGTGTGACCTCGACCGCGGGAATCGACGAAC 228
QY 146 IleLeuTrpHisAsnCysLeuIleGlyLysSerAspArgGlnLysLeuGlyGlnLys 165
DB 229 ATCTGTGGCATGAGTGGCCCATCGGCAGAGGAGCAGCAGGGTCTGCTGAACCAAGAG 288
QY 166 GlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAla 185
DB 289 GGCTCGCTCGTGTGGATCACTGGCTAAGCGGTTTCAAGGAAAGCAGCTCGCGTGC 348
QY 186 LeuSerArgGluLeuHisCysArgGlyHisIleuThrTyrValLeuAspGlyAspAsnLeu 205
DB 349 CTGAGCCCGAGCTGCACGCGCAGAGGCCACCTCAGCTACGCTCCTCGACGGGACAACTC 408
QY 206 ArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArg 225
DB 409 AGGCACGGCTGAACAGGGAACCTCAGCTTCGGAGCAGAGGACCGCGCGCAGACATCCGC 468
QY 226 ArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeu 245
DB 469 AGAGTAGGGGAAGTAGCAAGCTGTTCCGCCACGCTGGCCTCGTCTGATCGCCAGCCTC 528
QY 246 IleSerProTyrArgArgAspAlaCysArgAlaLeuLeuProHisSerAsnPhe 265
DB 529 ATATCGCCCTACAGAAGCAGACCGCGGTGTCCGATCTGCTGCCCAAGCACTCGTTT 588
QY 266 IleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeu 285
DB 589 ATCGAGGTGTTCTCTGGACGTCCGCTTCAAGTGTGGAAGCCAGGAGCCCAAGGCCTC 648
QY 286 TyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspProTyrGlu 305
```

```
Db 649 TACAGCTCGACGCGCGCAAAATCAAAGGTTTCCACGGCATCGACGATCTTACGAA 708
Qy 306 ProProIleAsnGlyGluIleValIleLysMetLysAspGluGluCysProSerProLys 325
Db 709 CGCGCGTCGGAGTGTGAGATAGTATGATCCATGTAAGTCCGGGACTGCGCTTGCCTGAA 768
Qy 326 AlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuGln 342
Db 769 TCGATGGCTGGTCACGTTGTGTCGTACCTTGAGACGAATGGTTTCTCTCCAG 819

RESULT 9
US-10-425-115-62889
; Sequence 62889, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Roba, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 62889
; LENGTH: 2611
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(2611)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_157357C.1
US-10-425-115-62889

Alignment Scores:
Pred. No.: 4,528-77 Length: 2611
Score: 896.50 Matches: 196
Percent Similarity: 63.54% Conservative: 34
Best Local Similarity: 54.14% Mismatches: 79
Query Match: 49.10% Indels: 54
DB: 20 Gaps: 7

US-10-829-432-4 (1-343) x US-10-425-115-62889 (1-2611)
Qy 7 AsnGlnThrGluProLeuValThrHisThrGlnProPro----- 20
Db 1151 AGCGCGACACGCGCGCGCACTGCCCATGACGACGCGCTCCCGGGGCTCTCGTGTCCCTG 1210
Qy 21 -----SerProAlaProGlyProLaserGlnGlnArgGlnGlnGlnGlnGlnGlnGln 38
Db 1211 TCCGCTCTCCGTTGCGGCTCCACGCAACGCGCGCGCGGGGTGGCGGTCTCTCGCG 1270
Qy 39 -----SerProThrProThrLeuAlaValIleLeuVal 49
Db 1271 ATCTCTGTGTGATTCACGCGCGGGGGTGGCGGTCATTCCTCCCTCGCGGCTTGCCTCGCG 1330
Qy 50 AsnProGlnArg-----AlaProValLeuProGlyLeuThrProSerAsp 65
Db 1331 TCGCGGTCCGCGCACCGGACCGGACCGCGCGCGGATCTA-GGGTGTGCTCCGCTCGC 1389
Qy 66 AlaProLeuProAlaValIleHisGlyLeuThrProArgSerHisSerAla 85
Db 1390 CCCCTCTCCGAGACTCTAGCGGCGAGGAATTAGGACGAGGTA----- 1434
Qy 86 GlyLeuAlaSerAspSerGlyArgGluGlyGluGlyArg----- 99
Db 1435 -----ACAGCGACACAGGCTCGCAAGGACGCGCCCTCGTCAATGCGCC 1479
Qy 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTyrValArgArgArgAsn 119
```

```
Db 1480 GCGCGCGCATCTC-----GTGGAGGAGCAGCGCGGAGCAC 1512
Qy 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
Db 1513 GCAGGAGTTGAAGGGCAAGCT-----TTGAAAATGTCTATCCACT 1551
Qy 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db 1552 GTGCGGAGTCAATCAATATCTTCTGGCATGATTCCTTGGCTAGTTGGCAAGCTGATGCCAG 1611
Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db 1612 AAGCTACTCAACACAGAAAGGCTGTCTGTCTGGATCACAGGCTTAGTGGCTCAGGTA 1671
Qy 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
Db 1672 AGTACCTTGGCATGTACATTAGGCGGGAGCTCCACACAGAGGGAAGCTTGCATATGTT 1731
Qy 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelysAlaGluAsp 219
Db 1732 CTTGACGGTGATTAAGACATGCTTGAACAGGATCTTGGCTTTAAGCTGAGAC 1791
Qy 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 1792 CGTCTGAAAATATAAGGAGAGAGTTGGTCAAGTAGCAAGTTATTTCAGATGCGGGCCTT 1851
Qy 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
Db 1852 GTATGTATTGCAAGTTTGTATATCTCATATAGGAGAGACCGTGAATCTTCCGCTGCATTG 1911
Qy 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 1912 TTGCTGTAGTAGTCTTCAATTGAAAGTTTCTTGAACATGCTCTTGAATATTGTGAAGCA 1971
Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db 1972 AGAGATCAAGGGGCTCTATAAGCTTCTCGTCAGGAGAAAATAAAGGTTTTCACAGA 2031
Qy 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db 2032 ATAGATGACCTTATGAAGACACCACTGAATTTGTGAGATTGAGATCAAGAGGTAGATGTT 2091
Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 2092 GTATGCCCTCCGCTCTGAGATGCGAGGCGCAAGTTGTTACTTACCTTGAGGAGAAAGCC 2151
Qy 340 TyrLeu 341
Db 2152 TTCCTG 2157

RESULT 10
US-10-425-114-18247
; Sequence 18247, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18247
; LENGTH: 1180
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
```

OTHER INFORMATION: Clone ID: LIB3079-005-C3_FLI
US-10-425-114-18247

```
Alignment Scores:
Pred. No.:      4,266-77      Length:      1180
Score:          892.50      Matches:      196
Percent Similarity: 65.32%      Conservative: 30
Best Local Similarity: 56.65%      Mismatches: 71
Query Match:      48.88%      Indels:      50
DB:              18          Gaps:      10

US-10-829-432-4 (1-343) x US-10-425-114-18247 (1-1180)

QY      7  AenGlnThrGluProLeuValThrHisThrGlnGlnProProSerProAlaProGlyPro 26
DB      6  ACCCAGACCGCTCGCGCTCTCCAC-----CGCCTCCAGGTACGCGCGCGCCCCC 59

QY      27 AlaSerGlnGlyGlnArgGlnGlyAsnThrLeuLeuSerProThrProThrLeuAlaVal 46
DB      60 TCCTCC-----CTCCTCCCTCCACCGCCA-----83

QY      47 IleLeuValAsnProGlnArgAlaProProValLeuProGlyLeuThrProSerAspAla 66
DB      84 -----CCCCAAAGCGCCACCG-----CGCCAGCTACTGCCTCTCCCGTA 125

QY      67 ProLeuProAlaLeuValIleHisGlyLeuThrProArgSerSerHisSerAlaGly 86
DB      126 CCG-----CGCGCCCATCGTACGACGAGCGCGGTAGGG 158

QY      87 LeuAlaSerAspSerGlyArgArgGlu-----GlyGluGlyArgGlyAla-----101
DB      159 GTGGGT-----GGTGCCGAAGGAGGAGGATGACGAGCGCGCGCGGTGATA 212

QY      102 ArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsnGlyAla 121
DB      213 CGGGCACACCTCTCTCTGACTTCAAGCGCGGTGCGTGC-AGGAAAAGTACTGCTGTC 271

QY      122 AlaPro-----GlyGluAlaProHisSerProValLysGluLysPro 135
DB      272 CCACCGCTCGCTGGNAAGTTGAAGGCAAGCT-----TTGAAA 310

QY      136 ValMetSerAsnIleGlySerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGln 155
DB      311 ATGTCTATCCACTGTGCGGAAGTCATCAATATCTTCTGGCATGATTCCTAGTTGGCAAG 370

QY      156 SerAspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSer 175
DB      371 ACTGATCGCCAGAAGCTACTCAACCAAGAAAGCTGTGTGTCTGGATCACAGGCTTAGT 430

QY      176 GlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHis 195
DB      431 GGCTCAGGTAAAGTACCTTGGCATGTACATTAGCGCGGAGCTCCACACAGAGGGAAG 490

QY      196 LeuThrTrpValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhe 215
DB      491 CTTCATATGTTCTTGACGCTGATAACTTAAGACATGTTCTGAAACAAGGATCTTGCTTT 550

QY      216 LysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAla 235
DB      551 AAAGCTGAAGATCGTGCTGAAATATTAAGGAGATGTTGGTGAAGTAGCAAAAGTTATTTCA 610

QY      236 AspAlaGlyValIleCysIleAlaSerLeuIleSerProTrpArgArgAspArgAspAla 255
DB      611 GACGCTGGCTTGATGATTGTCNAGTTTGATATCTCCATATAGGAGACCGTGAATCT 670

QY      256 CysArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLys 275
DB      671 TGCCGTGCACCTGTGTGCGGATGGTAGCTTTATTGAAGTTTCTTGAATATGTCCTGGAA 730

QY      276 IleCysGluAlaArgAspProLysGlyLeuTrpLysLeuAlaArgThrGlyLysIleLys 295
DB      731 TTATGTGAAGCAAGGAGCCCAAGGGCCCTTATAGCTTGCTTCGCTGGGAAATNAAG 790
```

```
QY      296 GlyPheThrGlyIleAspAspProTrpGluProProIleAsnGlyGluIleValIleLys 315
DB      791 GGTTTTACAGCAATAGATGACCTTATGAACACCACCTGAATTGTGAGATTGAGATCAAG 850

QY      316 MetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTrpLeu 335
DB      851 GAAATAGATGTTGATGCTTCCTTCGCTTCTGACATGGCGGGCAAGTGTGTACTTACCTT 910

QY      336 GluGluAsnGlyTrpLeu 341
DB      911 GAGGAGAAAGGCTTCCTG 928

RESULT 11
US-10-829-432-1
; Sequence 1, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; PRIOR FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Zea mays
US-10-829-432-1

Alignment Scores:
Pred. No.:      1,04e-76      Length:      890
Score:          887.00      Matches:      171
Percent Similarity: 84.23%      Conservative: 16
Best Local Similarity: 77.03%      Mismatches: 35
Query Match:      48.58%      Indels:      0
DB:              19          Gaps:      0

US-10-829-432-4 (1-343) x US-10-829-432-1 (1-890)

QY      121 AlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsnIle 140
DB      6  GCGCGCGCGCGCTCGCAGGAGTACGACGACGACGCGCTGTGACCTCGACCCGTC 65

QY      141 GlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLys 160
DB      66 GGGAAATCGACGAACATCTCTGTGCATGAGTGGCCATCGGCGCAGAGGACGACAGGT 125

QY      161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer 180
DB      126 CTGCTGAACCAAGAGGGCTGCGTCTGTGTGATCACTGGCCCTAAGCGGTTACGGGAAAAAGC 185

QY      181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpValLeu 200
DB      186 ACGTTCGGTGGCGGCTGAGCGGAGCTGACGCGCAGAGGCGCCTCACGTACGTCCTC 245

QY      201 AspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArg 220
DB      246 GACGCGACACCTCAGGCACGGCTGAACAGGACCTCAGCTTCGGAGCAGAGGACCGC 305

QY      221 AlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIle 240
DB      306 GCCGAGAAATCCCGCAGAGTAGGGGAAGTAGCAGAGCTGTTGCGCCGACGCTGGCCCTGTC 365

QY      241 CysIleAlaSerLeuIleSerProTrpArgArgAspArgAlaCysArgAlaLeuLeu 260
DB      366 TGCATCGCCAGCTCATATCGCCCTACAGAGCGACCGAGCGCGGTGTCGATCTGCTG 425
```



```
QY 21 -----SerProAlaProGlyProAlaSerGlnGlyGlnArgGlnGlnGlyAsnThrLeuLeu 38
Db 79 TCCGCTCCCGTTTCGGCGCTCCACAGCAACGCGCGCGGGTGGCGGTGCTCTCGCG 138
QY 39 -----SerProThrProThrLeuAlaValLeuVal 49
Db 139 ATCTCTGTCGATTCACAGCGCGGGGTGCCATTCCTCCCTCGCGGTTCGCCGCGCC 198
QY 50 AsnProGlnArg-----AlaProValLeuProGlyLeuThrProSerAsp 65
Db 199 TCGCGGTCCGCGACCGCGAACCGGAGCCCGCGCGGAATCGA-GGGTTGCTCCGCTCGC 257
QY 66 AlaProLeuProAlaLeuValLeuHisGlyLeuThrProArgSerSerHisSerAla 85
Db 258 CCCCTCTCCGAGACTCTAGCGCGAGGATTAGCAGCGGTA-----302
QY 86 GlyLeuAlaSerAspSerGlyArgGluGlyGluGlyArgGlyAlaArgThrHisCys 105
Db 303 -----ACAGGCGACAGAGGCTCGGAAGGACGCC-CTTCGTCAA-----340
QY 106 HisArgGlyIleGlyArgTrpValArgArgArgArgAenGlyAlaAlaProGlyGlu 125
Db 341 -----ATGCGCGCGCGACCGCATCCGTGGAGAGCAGCGGAGC-378
QY 126 AlaProHisSerProValysGluLysProValMet-----SerAsnIleGlyLys 142
Db 379 ACGCAGGAGGTAGTCTAGTTCGAAGCGCAAGCTTTGAAAATGTCATCCACTGTGCGGAAG 438
QY 143 SerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeu 162
Db 439 TCATCAAAATATCTTCGTGCATGATTCCTAGTTGGCAAGACTGATCGCCAGAAAGTACTC 498
QY 163 GlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyValSerThrLeu 182
Db 499 AACCAAGAGGCTGTGTCTCGGATCACAGGCCCTTAGTGGCTCAGGTAAGATACCTTG 558
QY 183 AlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGly 202
Db 559 GCATGTACATATTAGCGCGGAGCTCCACACAAGAGGGAAGCTTGCATATGTTCTTGACGGT 618
QY 203 AspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGlu 222
Db 619 GATAACTTTAAGACATGGTCTGAACAGGATCTTGCGCTTTAAAGCTGAAGACCGTGCATA 678
QY 223 AsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIle 242
Db 679 AATATAAGGAGAGTTGGTAGTAGCAAAAGTTATTTTCAGACCGCGGCTGTGTATGATT 738
QY 243 AlaSerLeuIleSerProTyrArgArgAspArgAlaCysArgAlaLeuLeuProHis 262
Db 739 GCAAGTTTGATATCTCCATATAGGAGAGACCGTGAATCTTGCGGTGCAATTGTTATCGGAT 798
QY 263 SerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspPro 282
Db 799 GGATGCTTTATTAAGAGTTTCTTGAATATGCTCTTGAATATTATGTAAGACGAAGGACCCA 858
QY 283 LysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAsp 302
Db 859 AAGGGCCTCTATAGCTTGCTGCTGCTGGGAAATAAAGGGTTTACAGGAATAGATGAC 918
QY 303 ProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGluCysPro 322
Db 919 CCTTATGAAGCACCACTGAATTTGTGAGATTGAGATCAAGGAATAATAGATGGTGTATGCCT 978
QY 323 SerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluLysAsnGlyTyrLeu 341
Db 979 TCGCCTTCTGACATGGCGGGCAAGTGGTTACTTACCTTGAAGGAAAGGCTTCCTG 1035
```

RESULT 14

US-10-425-115-135127

; Sequence 135127, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 135127
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_54717C.1
US-10-425-115-135127
```

Alignment Scores:

Pred. No.:	3.5e-73	Length:	1764
Score:	854.50	Matches:	203
Percent Similarity:	60.41%	Conservative:	32
Best Local Similarity:	52.19%	Mismatches:	83
Query Match:	46.80%	Indels:	73
DB:	20	Gaps:	10

US-10-829-432-4 (1-343) x US-10-425-115-135127 (1-1764)

```
QY 4 HisPheIleAsnGlnThrGluProLeuValThrHisThrGln-----17
Db 356 CATTTTGAAGTCAACTTGTCCTTTTAACTACTCATATAGATTTTAGTCACACTTCAAGT 415
QY 18 -----19
Db 416 TTTAAACATGTTTATTGTCAGGTCCGGATGGACGACGCTCCCGGCGGCCACCCCG 475
QY 20 ProSer-----ProAlaProGlyPro-----AlaSerGln 29
Db 476 CCCGCGTGGCGTGCCTGTCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 535
QY 30 Gly-----GlnArgGlnGlyAsnThrLeuLeu 38
Db 536 GGGGTGGTGGTCCCTCGCGTCTCTGTGTTCGATCCAGCGCGGAGGGTGCCTGTGCTT 595
QY 39 Ser-----ProThrProThr-LeuAlaValIleLeuValAsnProGlnArgAlaPr 55
Db 596 TGCTTCCCTGGCGTGGTCCCGCGCTTCGCGTCCGCGCACCTGAAC-----CGGAGGCC 649
QY 55 oProValLeuProGlyLeuThrProSerAspAlaProLeuProAlaLeuValIleHisG1 75
Db 650 CGCGTCAAAATCTAGGGTTGCATCCGCGCGGCCCTCTCTCCGAGACTC-----698
QY 75 yLeuThrProArgSerSerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgG1 95
Db 699 -----CCAGCGCTGGGAAC-CAGAACGAGGGGTGACGCGGACG 735
QY 95 uGlyGluGlyArgGlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgAr 115
Db 736 AAAGTCGACGAGCGGCCCTCTCGTCAATGCAC-----TCG-----CG 774
QY 115 gArgArgArgAsnGlyAlaAlaProGlyGluAlaProHisSerProValLysGluLys-- 134
Db 775 ACCGTTCGTTCGAGGAGCAGCGCGGAGCAGCGCAGGAGGTT-AGTTTAGTTGAGGAAAAAGC 833
QY 135 ----ProValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeu11 153
Db 834 TTTGAAATGTCATCCACCCTGCGGAGAGTCAATAATATCTCTGGCATGATTGTCAGT 893
QY 153 eGlyGlnSerAspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrG1 173
Db 894 TGGCAAGACTGATCGCCAGAGGCTACTCAAGCAGAAAGGCTGTGTGTTGCTGGATTACAGG 953
QY 173 yLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysAr 193
```



```
Db 954 CCTTAGTGGTTCAGGTAAGAGTACCTTGGCATGTACATTAGGCCGTGAGCTCCACACAAG 1013
Qy 193 gGlyHisLeuThrTyValLeuAspGlyValAspLeuArgHisGlyLeuAsnArgAspLe 213
Db 1014 AGGGAAGCTTGCATATGTCTTGACCGGTGATACCTTAAGACATGGTCTGAACAAGGATCT 1073
Qy 213 uSerPheLysAlaGluAspArgAlaGluAsnLeuArgValGlyGluValAlaLysLe 233
Db 1074 TGGCTTCAAGCTGAAGACCGTGTCTGAAATATAAGGAGAGTTGGTGAAGTAGCAAGATT 1133
Qy 233 uPheAlaAspAlaGlyValLeuLysCysLeuAlaSerLeuLeuSerProTyArgArgAspAr 253
Db 1134 ATTTCAGATGCGGGCTTGTATGTATTCGAAGTTTGATATCTCCATATAGGAGACCG 1193
Qy 253 gAspAlaCysArgAlaLeuLeuProHisSerAsnPheLeuGluValPheLeuAspLeuPr 273
Db 1194 TGAATCTTCCGTCGCAATTTGTCTGATAGTAGCTTCATGTGAAGTTTCTTGAACATGTC 1253
Qy 273 oLeuLysIleCysGluAlaArgAspProLysGlyLeuTyLysLeuAlaArgThrGlyLy 293
Db 1254 CTTGGAATTATGTGAAGCAAGAGATCCAAAGGGCTCTATAAGCTTGCTGTCGAGGAAA 1313
Qy 293 sIleLysGlyPheThrGlyLysAspProTyArgGluProProLysLeuGluValVa 313
Db 1314 AATAAAGGGTTTACAGGAATAGATGACCCCTTATGAAGCACCACCTGAATTGTGAGATTGA 1373
Qy 313 lIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGluValLeuCy 333
Db 1374 GATCAAGAGGTAGATGGTGTATGCGCTCCGCTCTGCTGAGATGGCAGGCAAGTTGTATAC 1433
Qy 333 sTyLysLeuGluAsnGlyTyLysLeu 341
Db 1434 TTACCTTGAGGAGAAAGGCTTCTCTG 1458
```

RESULT 15

```
US-10-425-114-9652
; Sequence 9652, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9652
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700856117_FLI
US-10-425-114-9652
```

Alignment Scores:

Pred. No.:	2,26e-73	Length:	1116
Score:	854.00	Matches:	155
Percent Similarity:	87.92%	Conservative:	27
Best Local Similarity:	74.88%	Mismatches:	25
Query Match:	46.77%	Indels:	0
DB:	18	Gaps:	0

US-10-829-432-4 (1-343) x US-10-425-114-9652 (1-1116)

```
Qy 137 MetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSer 156
Db 281 ATGGCAACTCTGAGCAATTCACATATATCTTTGGCAAGATTGTCAAGTAGGAGGCTT 340
```

```
Qy 157 AspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrrIleThrGlyLysSerGly 176
Db 341 GAAAGGCGAGAGCTACTTAAACCAAAAGGGATGTGTGTATGGATTACTGGACTCAGCGGA 400
Qy 177 SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu 196
Db 401 TCAGGAAAAAGCACATTGGCATGTTCCTCAAGCAGAGAACTGCATCTCAAGGGAAGTTA 460
Qy 197 ThrTyValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLys 216
Db 461 TCTTATGTCTTGTATGGAGATAACCTTCGGCATGGACTAAACCAAGGATCTTGGTTTCAA 520
Qy 217 AlaGluAspArgAlaGluAsnIleArgValGlyGluValAlaLysLeuPheAlaAsp 236
Db 521 GCCGAGAGATCGGCTGAAAATAATTCCGAGAACTGGAGAAAGTGGCAAAACTGTTCGTG 580
Qy 237 AlaGlyValIleCysIleAlaSerLeuIleSerProTyArgArgAspArgAlaCys 256
Db 581 GCTGGCTTAAATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 640
Qy 257 ArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysLe 276
Db 641 CGTGCCATGTGTACCTGATGCAAAATTTTATTGAGGTTTTTCATGAACATGCCTCTAGA 700
Qy 277 CysGluAlaArgAspProLysGlyLeuTyLysLeuAlaArgThrGlyLysIleLysGly 296
Db 701 TGTGAGGCGCAGATCCAAAGGGCTCTATTAAGCTTGTCTGCGGGGAAATCAAGGT 760
Qy 297 PheThrGlyIleAspAspProTyArgGluProIleAsnGlyGluIleValIleLysMet 316
Db 761 TTTACCGGCATTGATGATCCTTATGAACCCCAACCCCACTGTGAGATTGAAATTAAGCAG 820
Qy 317 LysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuLeuCysTyLys 336
Db 821 GAAATGGGGATTGTCCAACACCCCACTTTGATGGCTGGCAGCAAGTAGTTACTTACTTGG 880
Qy 337 GluAsnGlyTyLysLeuGlnAla 343
Db 881 AATAAAGGATTTCTTGAATCC 901
```

Search completed: September 9, 2005, 15:30:30
Job time : 674 secs

THIS PAGE BLANK (USPTO)